AN INTRODUCTION

TO THE

THEORY OF OPTICS

BY

ARTHUR SCHUSTER,

PH.D. (HEIDELBERG), Sc D (CANTAB), F R S

PROFESSOR OF PHYSICS AT THE UNIVERSITY OF MANCHESTER

LONDON

EDWARD ARNOLD

41 & 43, MADDOX STREET, BOND STREET, W

1904

THIS VOLUME

IS DEDICATED TO

JOHN WILLIAM STRUTT

BARON RAYLEIGH

O M, Sc D, F.R S

WHO BY HIS WRITINGS HAS ADDED CLEARNESS AND PRECISION

TO NEARLY ALL BRANCHES OF OPTICS

PREFACE.

THERE is at present no theory of Optics in the sense that the elastic solid theory was accepted fifty years ago We have abandoned that theory, and learned that the undulations of light are electromagnetic waves differing only in linear dimensions from the disturbances which are generated by oscillating electric currents or moving magnets. But so long as the character of the displacements which constitute the waves remains undefined we cannot pretend to have established a theory of light. This limitation of our knowledge, which in one sense is a retrogression from the philosophic standpoint of the founders of the undulatory theory, is not always sufficiently recognized and sometimes deliberately ignored. Those who believe in the possibility of a mechanical conception of the universe and are not willing to abandon the methods which from the time of Galileo and Newton have uniformly and exclusively led to success, must look with the gravest concern on a growing school of scientific thought which rests content with equations correctly representing numerical relationships between different phenomena, even though no precise meaning can be attached to the symbols used. The fact that this evasive school of philosophy has received some countenance from the writings of Heinrich Hertz renders it all the more necessary that it should be treated seriously and resisted strenuously

The equations which at present represent the electromagnetic theory of light have rendered excellent service, and we must look upon them as a framework into which a more complete theory

must necessarily fit, but they cannot be accepted as constituting in themselves a final theory of light.

The study of Physics must be based on a knowledge of Mechanics, and the problem of light will only be solved when we have discovered the mechanical properties of the æther While we are in ignorance on fundamental matters concerning the origin of electric and magnetic strains and stresses, it is necessary to introduce the theoretical study of light by a careful treatment of wave propagation through media the elastic properties of which are known A study of the theory of sound and of the old elastic solid theory of light must precede therefore the introduction of the electromagnetic equations

The present volume is divided into two parts, the first part includes those portions of the subject which may be treated without the help of the equations of dynamics, although a short discussion of the kinetics of wave motion is introduced at an early stage The mathematical treatment has been kept as simple as possible, elementary methods only being used I hope that rigidity of method is nowhere sacrificed thereby, while the advantage is gained that students obtain an insight into what is most essential in the theory of Interference and Diffraction, without introducing purely mathematical difficulties such as are involved in the use of Fresnel's integrals Even accurate numerical results may be obtained by a proper use of Fresnel's zones

The second part of the book is intended to serve as an introduction to the higher branches of the subject It has not been my object as regards this more advanced portion to write a treatise which shall be complete in itself, but rather to introduce the student to the writings of the original authorities As a teacher, I consider this to be the correct method, being convinced that students should be encouraged at an early stage to consult the literature of the subject It is a necessary consequence of the point of view adopted that the treatment is somewhat unequal Where the author has nothing to say which is novel, or may remove

obscurities, the best thing he can do, is to content himself with a short summary, referring the reader for details to the available sources of information. A more lengthy exposition is justified where a simplification or some new matter can be introduced. It may be mentioned in this connexion that as far as I know the consideration of absorptive regions of finite range of frequency in the theory of selective dispersion is new, and has not previously been published

I have purposely abstained from entering into details of methods of observation or instrumental appliances. These belong more properly to the courses of laboratory instruction.

I hope that the short biographical notices of deceased authors who have made important contributions to the science will be found to be of interest.

The greater part of this book was already in type when Lord Kelvin's *Baltimore Lectures* appeared, I was still able to add some references to these lectures, though not to the extent I should have wished. In some of the later chapters repeated reference is made to Drude's *Lehrbuch der Optik* Students who desire to pursue the subject further, should also have access to Mascart's *Optique* and Lord Rayleigh's *Collected Works* My own indebtedness to Lord Rayleigh's writings and personal inspiration is greater than can be acknowledged by mere references to his papers, and I am therefore glad to be allowed to dedicate this volume to him.

I am obliged to Prof. Wilberforce and Mr W H. Jackson for having looked through the proofs of the greater portion of the work, and favoured me with their corrections and suggestions. I have also to thank Mr J. E. Petavel for the very valuable help he gave me in drawing out the figures, and Mr H. E Wood for taking the photographs of interference effects which have been used in preparing the plates.

ARTHUR SCHUSTER.

August, 1904.

CONTENTS.

PART I.

CHAPTER I.

PERIODIC MOTION

CHAPTER II

KINEMATICS AND KINETICS OF WAVE MOTION

CHAPTER III

PRELIMINARY DISCUSSION OF THE NATURE OF LIGHT AND ITS PROPAGATION

CHAPTER IV.

THE INTERFERENCE OF LIGHT

CHAPTER V.

THE DIFFRACTION OF LIGHT

CHAPTER VI

DIFFRACTION GRATINGS

CHAPTER VII

THE THEORY OF OPTICAL INSTRUMENTS

CHAPTER VIII.

THE PROPAGATION OF LIGHT IN CRYSTALLINE MEDIA

CHAPTER IX

INTERFERENCE OF POLARIZED LIGHT.

PART II.

CHAPTER X.

THEORIES OF LIGHT

CHAPTER XI

DISPERSION AND ABSORPTION

CHAPTER XII

ROTATORY EFFECTS

CHAPTER XIII

TRANSMISSION OF ENERGY

CHAPTER XIV.

THE NATURE OF LIGHT

PART I.

CHAPTER I.

PERIODIC MOTION

1. **The Simple Periodic Motion.** A motion which is repeated at regular intervals of time is called a periodic motion. The simplest kind of periodic motion is that in which a particle moves in a straight line, in such a way that its distance, x, from a fixed centre satisfies the equation

$$x = a \sin \omega (t - \theta) \qquad (1),$$

where t is the time and a and ω are constants The equation shows that the particle continuously oscillates between two points which are at a distance a from the centre This distance is called the amplitude.

The velocity (u) of the particle which moves according to (1) is

$$u = a\omega \cos \omega (t - \theta) \qquad (2),$$

and the acceleration (f) is

$$f = -a\omega^2 \sin \omega (t - \theta) \qquad (3).$$

The particle passes through its central position ($x = 0$) when

$$t - \theta = m\pi/\omega,$$

m being an integer. The velocity of the particle is then ωa when m is even, and $-\omega a$ when m is odd Hence the velocity has its greatest value when $x = 0$, but may be positive or negative according as the particle passes through its central position from the negative or from the positive side.

If the time t is increased by $2\pi/\omega$, no change is made in the values of either x or u, so that after a time interval of $2\pi/\omega$ the position and state of motion are the same. The period τ is called the "*time of oscillation*," the "*periodic time*," or simply the "*period*" of the motion. Its relation to the constant ω is expressed by the equation.

$$\tau = 2\pi/\omega.$$

Equations (1) and (2) may take different forms by a change in th value of θ Thus by writing $\omega\theta_1 = \omega\theta + \frac{1}{2}\pi$, we obtain

$$x = a \cos \omega (t - \theta_1) \qquad (1a),$$

$$u = -\omega a \sin \omega (t - \theta_1) \qquad (2a).$$

When dealing with one particle only, so that the origin of time ma be chosen according to convenience, we may adopt the simpler forms (either (1) or (1a), obtained by making θ or θ_1 equal to zero

I proceed to show that equations (2) and (3) are necessary con sequences of (1)

In Fig 1 consider a point P moving uniformly in a circle of radiu

$$a = OA$$

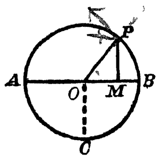

Fig 1

Let OM be the projection of OP on a diameter AB If the angle POM be denoted by ϕ, and the distanc OM by x,

$$x = a \cos \phi$$

If the particle passes through the position B whe $t = \theta_1$, and takes a time τ to complete a whole revolution,

$$\phi = 2\pi (t - \theta_1)/\tau.$$

Hence $$x = a \cos \omega (t - \theta_1),$$

where $$\omega = 2\pi/\tau$$

This shows that the point M moves in the simple periodic motio indicated by equations (1a) or (1) and we have the important pr position that this periodic motion may be represented as an orthogon projection of a uniform circular motion The periodic time τ is th time of revolution of the point P, the amplitude is the radius of th circle, and the constant θ_1 represents the smallest positive value of th time at which the particle reaches its extreme position on the positiv side

The proper expressions for the velocity and acceleration of th point M are obtained by considering that these are equal to the pr jections on AB of the velocity and acceleration of P.

The velocity of P is constant and equal to U

Hence $$u = -U \sin \phi$$
$$= -U \sin \omega (t - \theta_1)$$

The minus sign is explained by the fact that for positive values ϕ, the velocity of M is from right to left, or in the negative directio The whole circumference of the circle being described in a time t wi velocity U, it follows that

$$U = 2\pi a/\tau = a\omega.$$

Hence finally $$u = -a\omega \sin \omega (t - \theta_1).$$

The expression for the acceleration of the point M is obtained in a similar manner The acceleration of the point P is directed radially inwards towards the centre of the circle and is equal to U^2/a, and the acceleration f of M is the projection of this acceleration upon the diameter AOB.

$$\begin{aligned} f &= -(U^2 \cos\phi)/a \\ &= -a\omega^2 \cos\omega(t-\theta_1) \end{aligned}$$

A periodic motion may be of a more complicated character than that indicated by the above equations If we were to take *e g.* the orthogonal projection of a particle moving with uniform speed in an ellipse, we should get a motion which is strictly periodic, but which could not be represented by the simple equations we have given. Even the oscillations of a simple pendulum can only be *approximately* represented by our equations, the approximation being the more nearly correct, the smaller the amplitude.

I shall call a "simple" or "normal" oscillation one which can be represented as the orthogonal projection of a uniform circular motion. A normal oscillation is identical with that often called "harmonic motion." I avoid this term because "harmony" means a relation between different things, and not a property of any particular thing.

The character of the motion of a particle performing normal oscillations is completely determined by the amplitude and period, but the state of motion at any time requires a third quantity for its definition If the oscillation is considered to be the projection of a uniform circular motion, it is convenient to take the angle between the radius vector OP (Fig 1) and some fixed radius as the quantity defining the state of motion This angle is called the "phase" of motion, and is to a certain extent arbitrary, as the fixed radius may be drawn in any direction

If we express the motion in the form

$$x = a \sin\omega(t-\theta)$$

it is usual to define zero phase as the phase at the time the particle passes through its mean position in the positive direction The radius of reference will then be OC (Fig 1) at right angles to AB, and $\omega(t-\theta)$ will measure the phase

On the other hand, if we choose the form

$$x = a \cos\omega(t-\theta_1)$$

as the equation of motion, we may define zero phase as the phase at the time the particle reaches its extreme position on the positive side, then $\omega(t-\theta_1)$ will be the phase, and the radius of reference will be OB, or the positive branch of the direction on which the circular motion is projected The absence of uniformity in the choice of the direction

which defines the zero phase, causes no inconvenience, as we are nearly always concerned with *differences* of phase, and this difference is perfectly determinate. Thus if in Fig 6 two periodic motions are represented by the projections of the circular motions of two particles P and Q on the same straight line, the angle POQ will always represent the difference between the phases, whatever line is taken to be the direction of zero phase.

The difference in phase between two normal periodic motions having the same period is independent of the time

Representing the two motions by

$$x_1 = a_1 \cos \omega (t - \theta_1),$$

$$x_2 = a_2 \cos \omega (t - \theta_2),$$

the difference in phase will be

$$\omega (t - \theta_1) - \omega (t - \theta_2) = \omega (\theta_2 - \theta_1),$$

which proves the proposition

2. Normal Oscillations under the action of forces varying as the distance. The equations for the displacement x and the acceleration f of a particle which has a simple periodic motion are

$$x = a \sin \omega (t - \theta_1)$$

$$f = - a\omega^2 \sin \omega (t - \theta_1).$$

By combining these we obtain the relation.

$$f = - \omega^2 x \qquad (4)$$

This is an equation of great importance, for it shows the necessary condition which must be satisfied in order that a particle shall execute normal oscillations when acted on by a force directed to a centre. This condition is, that the force tending to bring the particle back to its position of equilibrium is proportional to the distance of the particle from that position

Consider a particle constrained to move in a straight line and attracted to a fixed centre by a force F, which is proportional to the displacement If m is the mass of the particle and $F = - n^2 x$

$$f = \frac{F}{m} = - \frac{n^2}{m} x.$$

This agrees with (4) if ω^2 is equal to n^2/m, and hence τ the time of oscillation is obtained in terms of m and n, for

$$\tau = \frac{2\pi}{\omega} = \frac{2\pi\sqrt{m}}{n}.$$

All forces of nature diminish with increasing distance, and the particular law of force which produces normal oscillations may not

first sight seem therefore to have any practical importance. As a matter of fact, this law, however, holds in almost all cases when the displacements are small, for when the particle is kept at rest under the action of opposing forces, the resultant of these forces will always, if the displacements are sufficiently small, increase proportionally to the distance of the particle from its position of equilibrium As this is an important fact, it is well to give a few examples

Example 1 *The Simple Pendulum.* A heavy particle is suspended from a fixed point by a thin string of length l and is set in motion.

Let θ (Fig 2) be the angular deviation of the string from the vertical at any instant The only forces acting on the particle are its weight and the tension of the string. The particle is constrained to move in a circle, and the force which tends to draw back the particle to its position of equilibrium is found by resolving these forces along the tangent to the arc.

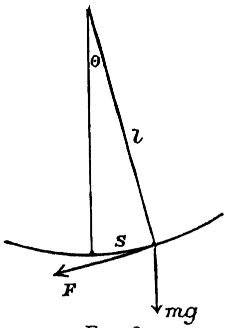

Fig 2

If m is the mass of the particle, its weight is mg. The tension of the string has no component in the direction of the tangent to the arc, and therefore the resultant force acting on the particle is $mg \sin\theta$

If θ is so small that we can neglect θ^2 compared to unity, we may replace $\sin\theta$ by the angle θ, so that the force F acting on the particle is

$$F = -mg\theta$$

$$= -mg\frac{s}{l} \qquad (5),$$

where s is the displacement of the particle along the arc corresponding to the angular displacement θ

This equation shows that the particle moves along the arc, as if it were subject to a restoring force which is proportional to the distance of the particle from the lowest point of the arc. Therefore the particle will describe normal oscillations about this point. The acceleration of the particle at any distance s is F/m or $-gs/l$ By comparing this with (4) it follows that $\omega^2 = g/l$ or that the period is determined by

$$\tau = 2\pi\sqrt{\frac{l}{g}}.$$

This is the well known equation for the time of oscillation of a simple pendulum

Example 2*. *Oscillations produced by volume elasticity of gaseous pressure.*

Let an airtight vessel be closed by a weighted piston A, which can move without friction in a cylindrical tube attached to the vessel This piston will have a definite position of equilibrium If it is forced down below this position and then released, it will be driven up again by the increased pressure of the air within the vessel The momentum it then acquires will carry it past its position of equilibrium The air in the vessel expands to fill the larger volume, its pressure is accordingly reduced, and it is unable to counterbalance the weight of the piston and the external pressure The piston is thus driven in again, and, the process repeating itself at regular intervals, periodic oscillations are performed.

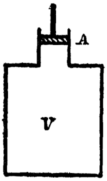

Fig. 3.

We proceed to find the time of oscillation of the piston Let V be the original volume of the vessel and P the pressure of the enclosed air Suppose the piston is pushed down until the volume is diminished by a small quantity v and the pressure is increased by a small amount p The volume and pressure are then $(V-v)$ and $(P+p)$ respectively. We shall disregard the inertia of the air and assume the motion to be sufficiently slow to allow the change to be isothermal We have then, applying Boyle's Law ·

$$VP=(V-v)(P+p)$$

or

$$pv=pV-vP$$

If the displacements are small, so that the product of the two small quantities p and v may be neglected,

$$pV=vP$$

or

$$p=P\frac{v}{V}$$

Denote by A the area of the end of the piston Then the resultant force on the piston, when the volume of the vessel is diminished by v, is

$$F=pA$$
$$=\frac{AP}{V}v$$

If x is the distance through which the piston moves

$$v=-A.x$$

and

$$F=-\frac{A^2P}{V}x$$

* Examples 2 and 3 are taken from Lord Rayleigh's *Sound* where they are treated in a different manner.

Thus the force acting on the piston is proportional to the displacement of the piston from its equilibrium position and is always in the opposite direction to the displacement. If M is the mass of the piston, its acceleration is F/M or $\frac{A^2P}{MV}\cdot x$ Hence the time of oscillation is given by

$$\tau = 2\pi\sqrt{\frac{MV}{A^2P}}$$

If the vessel were a cylinder of length l and area A, so that $V=Al$, and if also the pressure were entirely due to the weight Mg of the piston, we should have

$$P=\frac{Mg}{A},$$

and by substitution it would follow that

$$\tau=2\pi\sqrt{\frac{l}{g}},$$

i e the time of oscillation of the piston would be exactly the same as the time of oscillation of a simple pendulum, the length of which is the same as that of the cylindrical vessel

Example 3 *Normal Oscillations due to the tension of a string* Let a string attached to A pass over a peg B at the same level as A, and carry at its end a mass M If a particle P of mass m be attached to the string halfway between A and B, and the particle be displaced vertically downwards until it coincides with Q, the tension T of the string will have a resultant which, neglecting the weight of the string, is easily shown to be

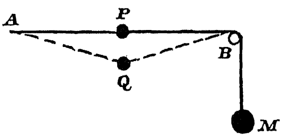

Fig 4

$$-\frac{2Tx}{\sqrt{a^2+x^2}},$$

where $2a$ is the distance AB and x the displacement PQ

If x is so small that x^2/a^2 may be neglected, and if m is so small compared to M that in the position of equilibrium the displacement of P is a small quantity of the second order, we may disregard the weight of P and take T to be equal to Mg The acceleration of P is therefore $-\frac{2Mgx}{am}$, and hence the time of oscillation

$$\tau=2\pi\sqrt{\frac{am}{2Mg}}.$$

3. Energy of a Particle in Periodic Motion. If a particle whose mass is m is moving with a velocity v, its kinetic energy is $\frac{1}{2}mv^2$

If the particle is executing normal oscillations, its velocity at any time t is

$$u = a\omega \cos \omega (t - \theta)$$

Therefore its kinetic energy E at this instant is

$$E = \tfrac{1}{2} m\omega^2 a^2 \cos^2 \omega (t - \theta)$$
$$= \tfrac{1}{4} m\omega^2 a^2 \{1 + \cos 2\omega (t - \theta)\}$$

The second term on the right-hand side has values ranging from $+1$ to -1, and is as often positive as negative, its average value taken over one complete period of vibration being zero

If $U = a\omega$ is the maximum velocity of the particle, it follows that the average value of E is $\tfrac{1}{4} m a^2 \omega^2$ or $\tfrac{1}{4} m U^2$ This proves that the *average* energy is half the *maximum* energy

The average value of the kinetic energy of a vibrating particle is taken as the measure of the intensity of the vibration, which has just been shown to be proportional to the square of the amplitude as long as the mass and period remain the same

Since no energy enters or leaves a vibrating particle, the sum of its kinetic energy and what is commonly called its potential energy must always remain constant. Now the kinetic energy E varies periodically, being at its maximum of $\tfrac{1}{2} m U^2$ when the particle is passing through its central position and falling to zero when the displacement is a maximum

If the constancy of the total energy is to be maintained, then it follows that the potential energy P must satisfy the equation

$$P + \tfrac{1}{2} m u^2 = \text{a constant}$$

Assuming the potential energy to be zero when the kinetic energy is at its maximum, the value of the constant must be $\tfrac{1}{2} m U^2$ Hence

$$P = \tfrac{1}{2} m (U^2 - u^2)$$
$$= \tfrac{1}{2} m a^2 \omega^2 \{1 - \cos^2 \omega (t - \theta)\}$$
$$= \tfrac{1}{2} m a^2 \omega^2 \sin^2 \omega (t - \theta)$$
$$= \tfrac{1}{2} m \omega^2 x^2$$

Thus for a body performing normal oscillations the potential energy is proportional to the square of the displacement of the particle from the centre of force

4. Composition of Periodic Motions. If a single particle is acted upon by two distinct agents, each of which, if acting separately, would cause the particle to perform simple periodic vibrations, the question arises—What is the resultant motion on the supposition that each produces its own effect?

We consider first the case in which the two component vibrations are in the same straight line and have the same period.

Let the two amplitudes be a_1 and a_2 and the common period, $2\pi/\omega$. At any instant t, the displacement x_1, due to the first oscillation, would be

$$x_1 = a_1 \cos \omega (t - \theta_1)$$

and that due to the second oscillation

$$x_2 = a_2 \cos \omega (t - \theta_2)$$

Since the two displacements are in the same straight line and each produces its own effect, we can combine them algebraically and write for the resultant displacement

$$\begin{aligned} x &= x_1 + x_2 \\ &= a_1 \cos \omega (t - \theta_1) + a_2 \cos \omega (t - \theta_2) \\ &= P \cos \omega t + Q \sin \omega t, \end{aligned}$$

where

$$\left.\begin{aligned} P &= a_1 \cos \omega\theta_1 + a_2 \cos \omega\theta_2 \\ Q &= a_1 \sin \omega\theta_1 + a_2 \sin \omega\theta_2 \end{aligned}\right\} \qquad (6)$$

Now write

$$P = R \cos \delta$$
$$Q = R \sin \delta,$$

so that

$$R^2 = P^2 + Q^2 \quad \text{and} \quad \tan \delta = Q/P$$

Then

$$\begin{aligned} x &= R (\cos \omega t \cos \delta + \sin \omega t \sin \delta) \\ &= R \cos (\omega t - \delta) \qquad (7) \end{aligned}$$

It is seen from the last equation that the two component simple periodic oscillations have combined to form a resultant simple periodic oscillation with the same period as the component oscillations, but with a different amplitude and phase

The amplitude of the resultant oscillation is R, where

$$\begin{aligned} R^2 &= P^2 + Q^2 \\ &= (a_1 \cos \omega\theta_1 + a_2 \cos \omega\theta_2)^2 + (a_1 \sin \omega\theta_1 + a_2 \sin \omega\theta_2)^2 \\ &= a_1^2 + a_2^2 + 2a_1 a_2 \cos \omega (\theta_2 - \theta_1) \end{aligned}$$

Therefore the amplitude $R = \sqrt{\{a_1^2 + a_2^2 + 2a_1 a_2 \cos \omega (\theta_2 - \theta_1)\}}$ is equal in magnitude to the diagonal of a parallelogram having two adjoining sides a_1 and a_2, the angle between the sides being $\omega (\theta_2 - \theta_1)$

We may now show that the diagonal OP not only represents the resultant oscillation in amplitude, but also indicates the phase.

By separating the quantities a_1 and a_2 in the equations (6) we obtain the two equations

$$\begin{aligned} P \sin \omega\theta_2 - Q \cos \omega\theta_2 &= a_1 \sin \omega (\theta_2 - \theta_1), \\ P \sin \omega\theta_1 - Q \cos \omega\theta_1 &= - a_2 \sin \omega (\theta_2 - \theta_1) \end{aligned}$$

If in these we substitute $P = R \cos \delta$ and $Q = R \sin \delta$, we find

$$\begin{aligned} R \sin (\omega\theta_2 - \delta) &= a_1 \sin \omega (\theta_2 - \theta_1), \\ R \sin (\omega\theta_1 - \delta) &= - a_2 \sin \omega (\theta_2 - \theta_1), \end{aligned}$$

or, expressed otherwise

$$a_1 \,.\, a_2 : R = \sin(\omega\theta_2 - \delta) : \sin(\delta - \omega\theta_1) : \sin\omega(\theta_2 - \theta_1) \qquad (8)$$

If a parallelogram be described with $OA = a_1$, and $OB = a_2$, as adjoining sides, and with an angle AOB equal to the difference in phase $\omega(\theta_2 - \theta_1)$, between the two oscillations which are to be combined, the geometry of Fig 5 shows that

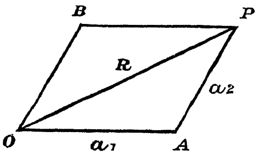

Fig 5

$$a_1 \;\; a_2 \;\; R = \sin BOP \;\; \sin AOP \,.\, \sin AOB$$

A comparison between this relation and (8) shows that

$$BOP = \omega\theta_2 - \delta \quad \text{and} \quad AOP = \delta - \omega\theta_1,$$

which means that the angles between the diagonal OP and the two lines a_1 and a_2 represent the difference in phase between the resultant oscillation and the two component oscillations

The proposition that normal oscillations in the same straight line may be combined like two forces is of primary importance, and a second proof of it is therefore given

We represent the two periodic motions by the orthogonal projections OM_1 and OM_2 of two points P and Q, moving with uniform speeds round two circles (Fig 6)

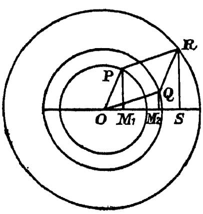

Fig 6

The two radii OP, OQ represent the amplitudes of the respective oscillations. If the periodic time is the same the radii OP, OQ revolve with the same angular velocity, and therefore the angle POQ remains constant Complete the parallelogram $OQRP$ and imagine a third point at the angle R Then the point at R will describe a circle in the same time as the points P and Q, and its projection S on the diameter AB will perform a simple periodic vibration

But $$OS = OM_1 + OM_2$$

since the projection of OR must equal the sum of the projections of OP and OQ.

Hence the displacement of S is always equal to the sum of the displacements of M_1 and M_2 and the motion of S will be the resultant of the motions of M_1 and M_2.

The figure shows that the resultant amplitude OR is found from the amplitudes OP, OQ by the parallelogram construction and that this construction enables us to determine not only the amplitude but also the phase of the resultant motion. For, if we measure phases from the time the particles have their maximum positive displacement,

then in the figure, the phase of the particle M_1 is the angle POS and the phases of M_2 and S are measured by the angles QOS and ROS respectively

Hence the direction of the diagonal OR indicates the phase of the resultant oscillation

5. Combination of any number of Oscillations. Having seen how two linear oscillations, which are in the same straight line, can be combined, it follows that any number of such oscillations can be combined by taking the resultant of any two of them, and combining with it a third oscillation and so on, until we reach the final resultant In short, a system of such oscillations is reduced to a single resultant in exactly the same way as a system of forces acting at a point. Any proposition relating to a system of forces can be made to apply to a system of linear oscillations of the same period which take place in the same straight line.

According to a well known proposition in Statics a system of n forces OP_1, OP_2 OP_n has a resultant which coincides in direction with OG and is in magnitude equal to nOG, if G is the centre of inertia of particles having equal masses, placed at points P_1, P_2 P_n We make use of this proposition to find the resultant of a large number of oscillations of equal amplitude and having their phases in arithmetic progression

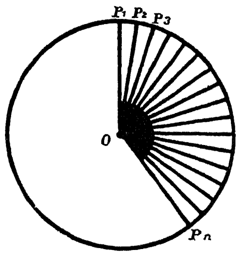

Fig. 7.

The oscillations will be represented by the lines OP_1, OP_2, OP_3, OP_n (Fig 7), such that all the points P are equidistant and lie on the arc of a circle If the constant phase difference between two successive oscillations is very small, the problem of finding the resultant resolves itself into the determination of G, the position of the centre of inertia of the arc P_1P_n

The distance OG is known to be equal to $\dfrac{a \sin \alpha}{\alpha}$ where 2α is the angle of the arc P_1P_n and hence α is the angle between OG and either OP_1 or OP_n The resultant vibration has therefore an amplitude equal to $na \sin \alpha/\alpha$ and a phase which lies halfway between the phases of the first and last vibrations If all vibrations were of equal phase, the resultant amplitude would be na

Hence we may formulate the following important proposition proved by the above reasoning

Normal rectilinear oscillations having equal amplitude and period, and taking place along the same straight line with differences of phase

such that any two successive oscillations have a phase difference which is small and equal for each successive pair, combine together into a resultant oscillation which has the same period, and a phase halfway between that of the first and last oscillation The amplitude of the resultant oscillation is $R \sin\alpha/\alpha$ where 2α is the phase difference between the two extreme oscillations and R the amplitude of the resultant in the special case that the phase differences vanish The values of $\sin\alpha/\alpha$ and $\sin^2\alpha/\alpha^2$ are plotted as ordinates as against α as abscissa in Fig 70, Art 54

6. Combination of oscillations in directions at right angles to each other. Let a particle M (Fig 8) describe simple periodic oscillations in the direction OX about the centre O, its motion being represented by the equation

$$x = a_1 \cos \omega t$$

Also let a second particle N perform oscillations of the same period, about the same centre O, but in the direction OY perpendicular to OX The motion of N may be represented by the equation

$$y = a_2 \cos(\omega t + \delta),$$

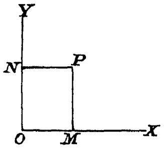

Fig 8

δ expressing the difference of phase between the two oscillations Now imagine a third particle P to move in such a way that its projections on OX and OY always coincide with the points M and N. The problem is to investigate the motion of the particle P. Before treating the question generally we may take a few cases, which are simple and of special importance

Case I Let $\delta = 0$ This means that both M and N pass through the centre O at the same instant, and that therefore the point P passes through O

The equations of motion of M and N are respectively

$$x = a_1 \cos \omega t,$$
$$y = a_2 \cos \omega t$$

By eliminating the time t from the equations, we obtain a relation between x and y, which determines the path described by P.

Thus $$\frac{x}{a_1} = \frac{y}{a_2} \quad \text{or} \quad y = \frac{a_2}{a_1} x$$

This is the equation of a straight line passing through the origin O. The cosines of the angles which OP forms with OX and OY respectively are $a_1/\sqrt{a_1^2 + a_2^2}$ and $a_2/\sqrt{a_1^2 + a_2^2}$ Projecting x and y on OP we see that the distance (r) of P from the origin is $r = \sqrt{a_1^2 + a_2^2} \cos \omega t$ Therefore the motion of the particle P is a simply periodic linear oscillation in the direction OP, having the same periodic time as its component vibrations, and an amplitude equal to $\sqrt{a_1^2 + a_2^2}$

Case II *Let* $\delta = \pm\frac{\pi}{2}$. This means that the particle M is passing through its mean position when N has its maximum displacement

The equations of motion are now

$$\left.\begin{aligned} x &= a_1 \cos \omega t \\ y &= a_2 \cos\left(\omega t \pm \frac{\pi}{2}\right) \end{aligned}\right\} \qquad (9)$$

$$= \mp a_2 \sin \omega t.$$

Eliminating t by squaring and adding the two equations, it is found that

$$\frac{x^2}{a_1{}^2} + \frac{y^2}{a_2{}^2} = 1.$$

Hence the path described by the particle P is an ellipse. In the special case of $a_1 = a_2$ the equation becomes that of a circle of radius a_1

The time occupied by P in moving round the ellipse or circle is the same as the periodic time of the linear vibrations. It is easily seen that if the phase of M moving along OX *exceeds* by a right angle the phase of N moving along OY, the motion will be from the positive axis of x to the positive axis of y, for according to equations (9) the particle P crosses the positive axis of x when $t = 0$ When $t = \pi/2\omega$ or after a quarter of a period, P is on the positive or negative branch of OY according as the upper or lower sign is taken Hence the positive sign in the second equation (9) indicates a clockwise and the negative sign an anti-clockwise revolution. The axes of the ellipse in which P moves are coincident with the axes of x and y

Case III or General Case Let δ now have any value whatever. The equations of motion are

$$\left.\begin{aligned} x &= a_1 \cos \omega t \\ y &= a_2 \cos (\omega t + \delta) \end{aligned}\right\} \qquad (10)$$

Then $$\frac{y}{a_2} = \cos \omega t \cos \delta - \sin \omega t \sin \delta$$

$$= \frac{x}{a_1} \cos \delta - \sin \omega t \sin \delta,$$

$$\therefore \sin \omega t \sin \delta = \frac{x}{a_1} \cos \delta - \frac{y}{a_2},$$

and from (10). $$\cos \omega t \sin \delta = \frac{x}{a_1} \sin \delta.$$

Squaring and adding we get

$$\sin^2 \delta = \frac{x^2}{a_1{}^2} - 2\frac{xy}{a_1 a_2} \cos \delta + \frac{y^2}{a_2{}^2} \qquad (11).$$

This is the equation of an ellipse the axes of which are not now, in general, parallel to the x and y axes (Fig 9). By putting $\delta=0$ or $\delta=\pm\frac{\pi}{2}$ in the general equation we return to the special cases I and II.

In the customary treatment of the kinematics of a particle, the point P is said to possess simultaneously velocities along OX and along OY which are respectively equal to the velocities of its projections Adopting this mode of expression, we may say that a particle having two simple periodic motions of equal period at right angles to each other, moves in general in an ellipse, the time of revolution being equal to the period of the oscillation. In special cases the ellipse may become a circle or a straight line.

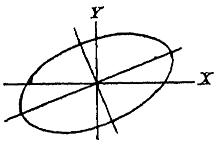

Fig. 9

The amplitudes a_1 and a_2 of two periodic motions at right angles to each other, and their difference of phase δ being given, we may require to determine the ratio of the principal axes of the elliptic orbit, and the inclination of these axes We obtain the result by means of some well-known propositions concerning curves of the second degree If a new system of coordinates be introduced inclined at an angle γ to the original one, so that x, y are related to x', y', by

$$\left.\begin{array}{l} x=x'\cos\gamma-y'\sin\gamma \\ y=x'\sin\gamma+y'\cos\gamma \end{array}\right\} \qquad (12),$$

the term in $x'y'$ disappears provided that

$$\tan 2\gamma=\frac{2a_1a_2}{a_1^2-a_2^2}\cos\delta,$$

or if

$$\tan\psi=\frac{a_2}{a_1},$$

$$\tan 2\gamma=\tan 2\psi\cos\delta \quad . \quad . \quad .(13)$$

The sum of the squares of the factors of x^2 and y^2 is not altered by the transformation If therefore A and B are the semi-axes of the ellipse, the substitution of (12) into (11) gives

$$\frac{1}{a_1^2}+\frac{1}{a_2^2}=\left(\frac{1}{A^2}+\frac{1}{B^2}\right)\sin^2\delta \qquad (14),$$

and similarly according to another well-known proposition, the product of the factors of x^2 and y^2 less the square of half the factor of xy remains constant in all transformations Hence

$$\frac{1}{a_1^2a_2^2}-\frac{\cos^2\delta}{a_1^2a_2^2}=\frac{\sin^4\delta}{A^2B^2},$$

$$\therefore \frac{1}{a_1a_2}=\frac{\sin\delta}{AB} \qquad (15).$$

Combining (14) and (15) we find

$$a_1^2 + a_2^2 = A^2 + B^2 \quad \ldots \quad (16),$$

and

$$\frac{a_1}{a_2} + \frac{a_2}{a_1} = \left(\frac{A}{B} + \frac{B}{A}\right) \sin \delta \quad \ldots \quad (17).$$

Writing $\tan \Psi$ for B/A, (17) gives

$$\sin 2\Psi = \sin 2\psi \sin \delta \quad \ldots \quad (18).$$

Eliminating δ between (18) and (13) we obtain

$$\cos 2\Psi \cos 2\gamma = \cos 2\psi \quad \ldots \quad (19).$$

Multiplying this last equation by (13) we find

$$\cos 2\Psi \sin 2\gamma = \sin 2\psi \cos \delta \quad \ldots \quad (20).$$

Finally the division of (18) by (20) gives

$$\tan 2\Psi = \sin 2\gamma \tan \delta \quad \ldots \quad (21).$$

We shall have occasion to use these formulae in Chapter XI

From (10) we obtain for the components of velocity u and v

$$u^2 = \omega^2 (a_1^2 - x^2),$$
$$v^2 = \omega^2 (a_2^2 - y^2)$$

Hence for the velocity in the elliptic orbit

$$U^2 = u^2 + v^2$$
$$= \omega^2 (a_1^2 + a_2^2 - r^2),$$

where r is the distance of the moving point from the origin We conclude that if the motion of a point in an elliptic orbit is capable of being represented by the superposition of two periodic motions at right angles to each other, the velocity in the orbit must follow a perfectly definite law If that law is satisfied the motion resolved along any axis is simply periodic.

A uniform circular motion, as we have seen, is equivalent to the superposition of two simple periodic motions at right angles to each other, the amplitudes being the same and the phase difference a right angle. The simple periodic motions may be taken along any two convenient directions at right angles to each other Conversely a simple periodic motion may be considered as being the superposition of two circular motions of equal periods and velocities in opposite directions. This may easily be proved either geometrically or algebraically.

Any number of simple periodic motions in a plane, having the same period but differing in amplitude and phase, may be combined into an elliptic motion This follows at once from the fact that each periodic oscillation may be decomposed into two along the same fixed axis at right angles to each other Adding the components which lie in the same direction according to Art. 4 and then combining the two

resultant oscillations at right angles to each other we obtain th required elliptic motion

Any number of simple periodic motions in a plane, having th same period but differing in amplitude and phase, may be combine into two uniform circular motions in opposite directions, but no necessarily along circles of equal radii This follows from the pro position that each of them may be decomposed into two opposit circular motions, together with the fact that two uniform circula motions in the same direction may be combined again into a uniforn circular motion

It follows that any elliptic motion in which the velocities satisf the required condition, may be considered as being the superpositio of two uniform circular motions in opposite directions.

To follow this out algebraically, let the rectangular projections of a circular motion taking place anti-clockwise be $a_1 \cos \omega t$ and $a_1 \sin \omega t$ anc that of another circular motion taking place clockwise $a_2 \cos \omega (t-\theta$ and $-a_2 \sin \omega (t-\theta)$ so that their combined motion is represented b

$$x = a_1 \cos \omega t + a_2 \cos \omega (t-\theta),$$
$$y = a_1 \sin \omega t - a_2 \sin \omega (t-\theta)$$

Eliminating t in the usual way, gives for the elliptic path th quadratic equation

$$x^2 (a_1^2 + a_2^2 - 2a_1 a_2 \cos \omega\theta) + y^2 (a_1^2 + a_2^2 + 2a_1 a_2 \cos \omega\theta) - 4xy a_1 a_2 \sin \omega\theta = (a_1^2 - a_2^2)^2$$

The three available constants a_1, a_2, and θ may now be determine in terms of the three constants which determine the elliptic orbit.

7. Composition of Linear Vibrations of slightly differen Periodic Times. We now consider the composition of two linea vibrations in the same direction but having slightly different periodi times

Let them be represented by

$$x_1 = a \cos \omega_1 t,$$
$$x_2 = a \cos \omega_2 t,$$

assuming, for simplicity, that they have the same amplitude Th resultant vibration is given by

$$\begin{aligned} x &= x_1 + x_2 \\ &= a \cos \omega_1 t + a \cos \omega_2 t \\ &= 2a \cos \frac{\omega_1 + \omega_2}{2} t \,.\, \cos \frac{\omega_1 - \omega_2}{2} t \end{aligned}$$

The factor $\cos \frac{\omega_1 - \omega_2}{2} t$ is periodic, varying in value between + 1

and -1 and going through a complete period in the time $4\pi/(\omega_1-\omega_2)$ Now this time is great (because $\omega_1-\omega_2$ is small) in comparison with the time $4\pi/(\omega_1+\omega_2)$ which is the period of the other factor We may therefore consider $2a\cos\frac{\omega_1-\omega_2}{2}t$ to be the slowly varying amplitude of a simple oscillation, having a period $4\pi/(\omega_1+\omega_2)$

The intensity I of the resultant vibration is proportional to the square of the amplitude, so that

$$I \propto 4a^2\cos^2\frac{\omega_1-\omega_2}{2}t$$

$$\propto 2a^2\{1+\cos(\omega_1-\omega_2)t\}.$$

Hence the resultant intensity varies between $4a^2$ and 0, and the time interval between two successive maxima of intensity is $2\pi/(\omega_1-\omega_2)$. An important application of this equation is made in the theory of sound When two notes of nearly equal pitch are sounded together, beats are heard, and according to the above, the periodicity of the beats is $2\pi/(\omega_1-\omega_2)$, if $2\pi/\omega_1$ and $2\pi/\omega_2$ are the periods of the two notes As the number of vibrations per second (the frequencies) are inversely proportional to the periods, it follows that when two notes have frequencies n_1 and n_2, the number of beats per second is n_1-n_2

8. Use of imaginary quantities. The mathematical treatment of oscillations may often be made more concise by the introduction of imaginary quantities Writing $\imath=\sqrt{-1}$, we make use of the symbolic expression

$$e^{\imath\phi}=\cos\phi+\imath\sin\phi$$

If $\phi=\omega t$, it is seen that both the real and imaginary part of $e^{\imath\phi}$, represents a simple periodic motion The same is true for $ce^{\imath\phi}$, where the "amplitude" c, may be real, imaginary, or complex. Writing, to separate the real and imaginary parts, $c=a+\imath b$, and

$$\left.\begin{aligned} r\cos\delta &= a\\ r\sin\delta &= b\end{aligned}\right\} \qquad\qquad (22),$$

it follows that $\qquad c=re^{\imath\delta}$ and $ce^{\imath\phi}=re^{\imath(\phi+\delta)}$

It follows from (22) that the amplitude r is equal to $\sqrt{a^2+b^2}$ and the phase δ determined by

$$\tan\delta=b/a$$

If the factor of $e^{\imath\phi}$ is of the form

$$c=\frac{a+\imath b}{A+\imath B}$$

The fraction is reduced to the standard form by multiplying its numerator and denominator by $A-\imath B$.

We derive $$c=\frac{(aA+bB)+i(bA-aB)}{A^2+B^2},$$

and hence the amplitude and phase of the real part of $ce^{i\phi}$ are obtained from

$$\left.\begin{aligned} r^2 &= \frac{a^2+b^2}{A^2+B^2} \\ \tan\delta &= \frac{bA-aB}{aA+bB} \end{aligned}\right\} \qquad (23)$$

For the particular case that $A=a$ and $B=-b$,

$$c=\frac{(a^2-b^2)+2iab}{a^2+b^2},$$

and $$r=1, \quad \tan\delta=\frac{2ab}{a^2-b^2}$$

According to the above, an expression of the form $\frac{a+ib}{A+iB}$ can always be brought to the form $re^{i\delta}$ where r and δ satisfy equations (23). We may write these equations in another way which is sometimes convenient Put

$$P=\frac{a+ib}{A+iB}, \qquad Q=\frac{a-ib}{A-iB},$$

where Q is obtained from P by reversing the sign of i It is then easily shown that

$$PQ=r^2, \quad \text{and} \quad \frac{P-Q}{i(P+Q)}=\tan\delta \qquad (24)$$

CHAPTER II

KINEMATICS AND KINETICS OF WAVE MOTION

9. Kinematics of Wave Motion. Every one is familiar with the appearance of a train of waves propagated over a surface of water. As a rule, such surface waves alter their shape as they proceed and they are not therefore very good examples of simple wave propagation. We say that a wave has "constant type" when the outline of the wave always remains the same Waves of sound and waves of light propagated through a vacuum are waves of constant type

Consider a row of particles lying on a straight line, which we shall take to be the axis of x Let the particles be displaced in a direction at right angles to x, the displacement being represented by the equation $y=f(x)$

If the displacements at each point alter in such a way that a line joining all particles seems to travel with a velocity v in the positive direction without change of shape, the equation of the outline at any time t may still be represented by the same equation $y=f(x)$, provided the origin from which x is measured is shifted through a distance vt Referred to the old origin, the equation representing the outline will be given by $y=f(x-vt)$ This then is the general equation of a wave of constant type propagated in the positive direction with a velocity v, and every wave propagated without change of shape must be expressible in this form The argument does not turn upon the displacement y being necessarily at right angles to the direction of propagation, but it may be, as in the case of sound waves, along that direction, and the equation would hold equally for displacements of any kind By giving to v a negative sign, we obtain the general equation of a wave propagated along x in the negative direction

Fig 10

As an example we may consider the equation

$$y = ae^{-(x-vt)^2}$$

which being of the form $y=f(x-vt)$ represents a wave motion

Putting $t=0$ we obtain for the shape of the wave, the outline

$$y = ae^{-x^2}.$$

The equation represents therefore a wave of the form shown in Fig 1 propagated with a constant velocity v in the positive direction

Returning to the general equation

$$y = f(x - vt)$$

we obtain by differentiation

$$\frac{dy}{dt} = -vf',$$

$$\frac{dy}{dx} = f',$$

$$\frac{dy}{dt} = -v\frac{dy}{dx} \qquad \text{(1)}$$

Also

$$\frac{d^2y}{dt^2} = v^2 f'',$$

$$\frac{d^2y}{dx^2} = f'';$$

$$\frac{d^2y}{dt^2} = v^2\frac{d^2y}{dx^2} \qquad \text{(2).}$$

The last equation is the differential equation which characteris a wave motion Its complete solution is

$$y = f(x - vt) + F(x + vt)$$

where f and F are arbitrary functions.

As an important special case we take

$$y = a\cos(\omega t - px) \qquad \text{(3).}$$

By comparison with the general expression, it is seen that ω/p the velocity of propagation If y is measured at right angles to x, a if each point always keeps the same distance from the plane $x=0$, t motion will be rectilinear For a given x the equation is of the form

$$y = a\cos(\omega t + \delta)$$

and every point therefore performs normal oscillations having a perio

$$2\pi/\omega$$

The outline of the wave is obtained by taking any value of *e.g.* $t=0$, when

$$y = a\cos px$$

will represent the shape of the wave and its position at that tim A portion of the wave form which reaches out to infinity in bo directions, is represented in Fig 11 The figure illustrates the meth of drawing the curve Equidistant points divide the circumference a circle into equal portions In the figure that number is twelve, b

could be increased if it is desired to obtain a greater number of points in the curve. Other equidistant points are taken on a straight line

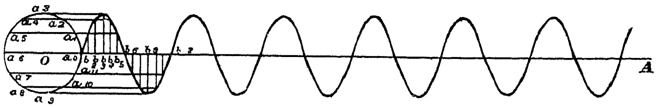

Fig 11

OA passing through the centre of the circle. Drawing perpendiculars to OA through each point on that line, and lines parallel to OA through the corresponding points of the circle, the intersections of the two sets of lines mark the points on the curve The *wave-length* is the distance between the two nearest points which have the same phase. If λ be the wave-length, so that the phase is the same for x and $x+\lambda$, it follows that $p\lambda$ must be equal to 2π, or $p=2\pi/\lambda$

From $v=\omega/p$ and $\omega=2\pi/\tau$, we obtain $v=\lambda/\tau$.

In terms of λ and τ we may write equation (3)

$$y=a\cos 2\pi\left(\frac{t}{\tau}-\frac{x}{\lambda}\right).$$

The difference of phase between two particles at distances x_1 and x_2 from the origin, as obtained from this equation, is

$$\frac{2\pi}{\lambda}(x_2-x_1).$$

In the further consideration of wave motion, we shall consider principally waves the displacement of which can be represented by the equation (3)

10. Application of Fourier's Theorem By an important theorem due to Fourier, any function $f(x)$ may between fixed limits $x=-c$ and $x=+c$ be represented as the sum of a series, in such a way that writing $\alpha=\pi x/c$

$$f(x)=a_0+a_1\cos\alpha+a_2\cos 2\alpha+a_3\cos 3\alpha+$$
$$+b_1\sin\alpha+b_2\sin 2\alpha+b_3\sin 3\alpha+\dots \qquad \dots \qquad . \; . \; .(4).$$

The constants a_0, a_1, b_1, b_2, etc may be determined from the function f, and we may for our present purpose fix for $f(x)$ outside the specified limits the values calculated from the series on the right-hand side If waves of all lengths are propagated with the same velocity v, we may obtain the shape at any subsequent time for waves travelling in the positive direction by writing in all terms on the right-hand side $x-vt$ for x, and having done so we may add the series again, when it is seen that the sum now becomes $f(x-vt)$ Hence the condition that normal waves of all lengths travel with the same velocity carries with it the consequence that waves of any shape may be propagated

without change of type On the other hand, if as in the case of light-waves travelling through a dispersive medium, the velocity of propagation depends on the wave-length, there must always be a change of type when waves which are not of the simple cosine or sine shape are propagated

11. Waves travelling along a stretched string. Let us now consider the kinetics of wave propagation

Consider a small portion AB of a curved string and acted on by equal tangential forces at the ends The resultant force passes by symmetry through C the centre of the circle of curvature of AB and bisects AB If 2θ be the angle subtended by the portion AB of the string at C, the intensity of the resultant is $2T\sin\theta$, which is nearly equal to $2T\theta$ if θ be sufficiently small As $2r\theta$ is the length of the arc AB, where r is the radius of curvature, the "resultant force per unit length" is T/r, *i e* equal to the product of the tension and the curvature

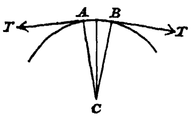

Fig 12

Let now a string be only slightly curved, so that every part of it is near a straight line which shall be the axis of x Its inclination to that axis dy/dx may be supposed to be sufficiently small to allow its square to be neglected The force acting on an element of length ds has been proved to be Tds/r, and neglecting the square of dy/dx, we may take the same expression to represent that component of the force which lies in the y direction

If ρ be the density per unit length, and hence ρds the mass of the length ds, the equation of motion is

$$\rho ds \frac{d^2y}{dt^2} = \frac{Tds}{r},$$

$$\therefore \frac{d^2y}{dt^2} = \frac{T}{\rho}\frac{1}{r}$$

Again neglecting $\left(\frac{dy}{dx}\right)^2$, the curvature is equal to $\frac{d^2y}{dx^2}$, hence

$$\frac{d^2y}{dt^2} = \frac{T}{\rho}\frac{d^2y}{dx^2} \qquad (5)$$

Comparing this equation with (2) it is seen that $\sqrt{T/\rho}$ is the velocity of the waves which are propagated along the string

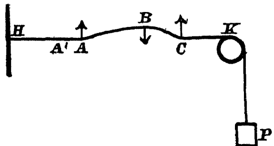

Fig 13

Let Fig 13 represent a portion of a string stretched to a constant tension by a weight P Let it be displaced by outside forces until it occupies a position such as that shown in the figure If the constraint is suddenly removed, the tension of the string will, by what precedes,

act in such a way that there is at each point a resultant force towards the centre of curvature. Hence the point B will begin to move downwards while A and C move upwards. If AH has been previously straight, this portion of the string is in equilibrium, but as soon as A is lifted up, the point at which the straight and curved portions join, has been moved to the left. If A' is that point, AA' which was previously in equilibrium, has ceased to be so It follows that a disturbance will set out from A and travel from right to left, with a velocity which has already been found to be $\sqrt{T/\rho}$ A similar reasoning shows that the displaced region AC will also send out a disturbance from C towards K Two waves travelling in opposite directions will therefore start from ABC

Now we know from observation that it is possible for a disturbance to travel in one direction only, and it is a matter of interest to examine the conditions under which a displacement such as ABC may be propagated forward only or backward only. In order that it shall travel only forward, it is clearly necessary that the point A should remain in its position in spite of the force acting upwards, and this is only possible if at the time to which the figure applies, A has a velocity downwards, of such magnitude that the force acting at A just destroys the velocity The force is of the nature of an "impulse" because if there is a discontinuity of slope at A, the curvature is infinite, and hence the force is infinite, and capable of suddenly destroying a finite velocity Similarly all along ABC a certain relation between velocity and displacement must hold, and this relation must be of such a nature that each portion will have zero velocity as soon as the wave has passed over it The mathematical relation which must connect the displacement and the velocity at each point when waves are propagated in one direction only, is obtained from (1) substituting the value of v

$$\frac{dy}{dt}\Big/\frac{dy}{dx} = \mp\sqrt{T/\rho},$$

where the upper sign holds for waves propagated in the positive direction.

I have discussed this question at length, because it shows clearly the important fact that if waves are sent out from any disturbed region, the displacements in that region are not by themselves sufficient to determine the subsequent motion, the velocities being just as important as the displacements In the above case, with the same displacements, the velocities might be chosen so as to give a wave wholly moving forward in one direction, or wholly moving back in the opposite direction, while generally there are two portions of the wave, one moving towards the positive, and one towards the negative side.

12. Transverse Waves in an Elastic Medium. We confine our attention for the present to bodies, the elastic properties of which are independent of direction Such bodies are said to be "isotropic."*

Consider a medium in which the displacements are the same in magnitude and direction for all points lying in the same plane drawn normally to a given line In Fig. 14 OX represents this line, and A_1B_1, A_2B_2, A_3B_3 are the intersections of a number of planes perpendicular to OX with the surface of the paper At each point of these planes the displacements are supposed to be identical, but they may differ in different planes If the displacements are all normal to OX and in the plane of the paper, each plane may be imagined to slide along itself through distances equal respectively to C_1C_1', C_2C_2' etc We confine the investigation to the case of elastic forces which are such that for the linear displacements contemplated, the restitutional force acts backwards in the direction of the displacement

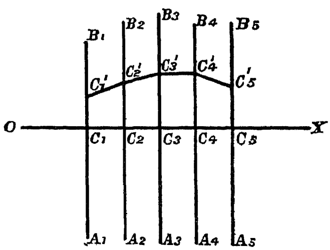

Fig 14.

The strain set up in the medium by the displacement is one involving change of shape only, and not any change of volume

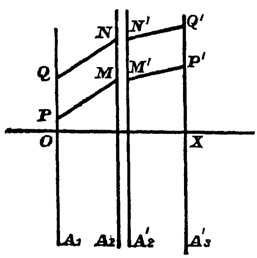

Fig 15

If PM and $M'P'$ are the positions in the strained condition of two lines originally parallel to OX, the parallelogram $PQNM$ was originally a rectangle, and the elementary theory connecting strains and stresses shows that the plane A_2B_2 to be maintained in its displaced position must be acted on by an upward force which per unit surface is equal to $n \tan \alpha$, where n is the resistance to distortion and α the angle between PM and OX Similarly the plane $A_2'B_2'$ to keep its position must be acted on by a downward force which per

* Thomson and Tait, Vol I, Art 676, give the following definition of isotropy:

"The substance of a homogeneous solid is called isotropic when a spherical portion of it, tested by any physical agency, exhibits no difference in quality however it is turned Or, which amounts to the same, a cubical portion cut from any position in an isotropic body exhibits the same qualities relatively to each pair of parallel faces Or two equal and similar portions cut from any positions in the body, not subject to the condition of parallelism, are undistinguishable from one another A substance which is not isotropic, but exhibits differences of quality in different directions, is called eolotropic"

unit surface is $n \tan \alpha'$, where α' is the angle between $P'M'$ and OX The resultant force acting on a small rectangular volume of unit height, thickness NN' and length MN is

$$MN \times n\,(\tan \alpha' - \tan \alpha)$$

If the displacements are denoted by y, we have

$$\tan \alpha = \frac{dy}{dx},$$

and if $MM' = t$,

$$\tan \alpha' = \frac{dy}{dx} + t\frac{d}{dx}\left(\frac{dy}{dx}\right),$$

so that the resultant force may now be written $MN \times t \times n\dfrac{d^2y}{dx^2}$, but $MN \times t$ is the volume considered, and if ρ is the density, $n\dfrac{d^2y}{dx^2}\Big/\rho$ will denote the resultant force divided by the mass We have considered the force necessary to maintain the medium in its strained condition, but if that force is removed, the acceleration may be obtained by the third law of motion:

$$\frac{d^2y}{dt^2} = \frac{n}{\rho}\frac{d^2y}{dx^2}$$

This equation is of the form (2) and shows that the medium is capable of transmitting waves in a direction OX with a velocity $\sqrt{n/\rho}$ As the velocity is independent of the wave-length, waves of any shape are propagated without change of type

If we imagine a second disturbance superposed on the one which has been discussed, and at right angles to it, we arrive at a wave propagation in which each particle describes a plane curve. We may for convenience limit the discussion to waves of the normal type, in which the displacements are therefore represented by

$$y = a\cos(\omega t - px)$$

Superposing a similar wave, the displacements being in the z direction,

$$z = b\cos(\omega t - px + \delta)$$

The paths of the particles in each plane are seen to be similar and elliptic, circular or rectilinear, according to the value of δ and the relations holding between a and b, Art 6

One important observation remains to be made Imagine the medium to consist of a number of detached particles, not acting on each other, but attracted to their position of equilibrium by a force varying as the distance

Let the particles at the time $t = 0$ be put into such a position and have such velocities that their displacements may be represented by

$$y = a\cos px$$

and their velocities by

$$v = -\omega a \sin px,$$

then the particles will continue to move under the action of the forces in such a way that their position at any subsequent time is represented by

$$y = a \cos(\omega t - px),$$

for this is the only relation which satisfies the condition that the accelerations are proportional to the displacements, and gives the required values for the displacements and velocities when $t = 0$ Hence a number of detached particles may simulate a wave motion, if once their displacements and velocities are properly adjusted, and if the force tending to bring them back to their position of rest causes an acceleration proportional to the displacement

13. Condensational Waves. We imagine the same conditions to hold as in the previous paragraph, with the exception that the displacement (ξ) shall be in the direction of propagation An investigation very similar to the one which was applied to the distortional or transverse waves will now hold, and it is not necessary to deduce again in detail the equation of motion, which for the case that $\frac{d\xi}{dx}$ is small is found to be

$$\frac{d^2\xi}{dt^2} = \frac{m}{\rho}\frac{d^2\xi}{dx^2}$$

Here m represents the longitudinal stress per unit elongation It would be wrong to substitute for m the resistance to dilatation, or, as one might be tempted to do, Young's Modulus The magnitude of m in terms of the elastic constants needs to be specially determined by the fact that there are no displacements at right angles to the direction of propagation This we proceed to do If the forces acting in the medium were all in the direction OX, a contraction of the medium at right angles to the direction of propagation would take place The application of Young's Modulus would be justified in that case, but we have worked under the assumption that the displacements (*not* the forces) are parallel to OX To counterbalance the contraction, transverse forces must act, and these forces will affect the elongations It is known from the elementary theory that if P be the normal tensional force along OX, it will produce an elongation equal to

$$P\left(\frac{1}{9k} + \frac{1}{3n}\right)$$

where k is the resistance to compression, and n the resistance to distortion.

The contraction at right angles is

$$P\left(\frac{1}{6n} - \frac{1}{9k}\right)$$

If equal tensions Q act along OY and OZ at right angles to OX, the elongations along OY and OZ are both equal to

$$Q\left(\frac{1}{9k}+\frac{1}{3n}\right)-(P+Q)\left(\frac{1}{6n}-\frac{1}{9k}\right)=Q\left(\frac{2}{9k}+\frac{1}{6n}\right)-P\left(\frac{1}{6n}-\frac{1}{9k}\right) \quad \ldots \quad (6)$$

The elongation along OX is, taking account of Q,

$$P\left(\frac{1}{3n}+\frac{1}{9k}\right)-2Q\left(\frac{1}{6n}-\frac{1}{9k}\right)$$

Substituting the value of Q found by equating (6) to zero, the elongation becomes

$$\frac{3P}{3k+4n}$$

The stress per unit elongation is therefore

$$m=k+\tfrac{4}{3}n$$

The velocity of propagation is $\sqrt{m/\rho}$ and depends therefore on the resistance to distortion, as well as on the resistance to compression.

The waves which involve longitudinal displacements only, are called condensational waves, because they involve changes of volume, but all condensational waves involve also distortion A difficulty may be found in admitting the existence of waves having the type considered on account of the force Q which would have to be applied at the boundary of the medium The difficulty no doubt exists in some cases and it would be wrong, for instance, to apply the result obtained, to the propagation of waves along a rod or bar Waves in which the displacements are solely in the direction of propagation could not travel along a rod, unless forces were applied at the surface and adjusted so as to prevent all contraction or expansion at right angles to the rod

In an elastic medium, the boundaries of which are at a considerable distance, plane waves do not occur except as the limiting case of spherical waves, when the radius of the sphere has become very large. There is no difficulty in conceiving radial displacements and stresses across planes of AB and $A'B'$ (Fig. 16), which prevent the lateral contraction. Our investigation may therefore be considered to apply to such spherical waves having a large radius

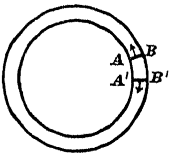

Fig 16

14. Spherical Waves If a disturbance is produced within a small volume of an isotropic elastic medium, it spreads out in the form of spherical waves Let at any one time, a very small volume T be disturbed, the rest of the medium being in a state of equilibrium. If all disturbing forces are now removed from the region T, the complete

theory proves, what the results of the previous paragraphs already lead us to expect, that v being the velocity of propagation, the disturbance after a time t, will be confined to the neighbourhood of a spherical surface drawn with uniformly increasing radius $r = vt$ about some point within T If the medium can propagate both distortional and condensational waves, the disturbance in general separates into two portions, one of these is spread over a sphere of radius $r_1 = v_1 t$, and consists of displacements which do not involve any change of volume, while the other, spread over the sphere of radius $r_2 = v_2 t$, involves both condensation and distortion In terms of the elastic constants, the velocities of propagation are the same as for plane waves, so that

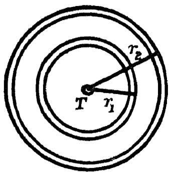

Fig 17

$$v_1 = \sqrt{n/\rho}, \qquad v_2 = \sqrt{(k + \tfrac{4}{3}n)/\rho} \quad . \qquad (7)$$

In all fluid media, the resistance to change of shape is zero, hence the distortional wave does not exist, and the condensational wave is propagated with velocity $\sqrt{k/\rho}$, where for rapid oscillations, such as take place in sound waves, k is the adiabatic and not the isothermal elasticity. If a medium is incompressible, k is infinitely large, and the condensational wave is propagated with infinite velocity

If the disturbance is of the normal periodic type, waves spread outward from the source, and, in consequence, energy is propagated outwards Unless there is a continuous accumulation of energy in space, the energy passing in unit time through all closed surfaces surrounding the centre of disturbance, must be the same Apply this to spheres of different radii drawn round the centre, when it will be clear that as the total energy transmitted through each sphere is the same, the energy per unit surface must be inversely proportional to the square of the distance.

Remembering (Art 3) that the energy of a particle performing periodic oscillations is proportional to the square of the amplitude and following the analogy of plane waves, we are tempted to write for the displacements (y) in a spherical wave,

$$y = \frac{a}{r} \cos 2\pi \left(\frac{t}{\tau} - \frac{r}{\lambda}\right) \quad \ldots \quad . \qquad (8),$$

where a is a constant which may be different for different directions, but remains the same along the same radius This is not, however, the correct expression (Chapter XIII) though it is approximately accurate, when r is large compared to λ, and becomes more and more nearly true in proportion as $\lambda/2\pi r$ is negligible

According to (8) the difference in phase between two points at a distance $r_2 - r_1$ from each other, along the same radius, would be $2\pi(r_2 - r_1)$ but this result is limited to the same restrictions as the equation itself and must not be applied when r_1 is not large compared with λ. Difficulties which have been felt in certain parts of the subject are due to the tacit assumption that (8) is generally correct, and that the difference in phase between a point at a distance r from the origin and a point at the origin is $2\pi r/\lambda$ This is not true

15. Waves spreading from a disturbed region of finite size. If the original disturbance be spread over a space T of finite dimensions, Fig 18, we may by a simple geometrical construction find the space which at any subsequent time t may be disturbed in consequence of the wave motion spreading out from T We assume that no forces continue to act within T, that space being left to regain a state of equilibrium under the action of its own elastic forces only.

Subdivide T into indefinitely small portions and consider each small portion to be an independent centre of disturbance, from which spherical waves spread out as in the last paragraph If A and B are the two points in T which are nearest and furthest, respectively, from P, then at a time AP/v the disturbance from A has just reached P Previous to that time the point P was at rest It will continue to be traversed by waves coming from some point in T until a time equal to BP/v Then the disturbance will have completely passed over it, and P will again be in equilibrium, *i e* its velocity will remain zero, though its position may be different from that which it occupied previous to the passage of the wave To obtain the region over which the disturbance is spread at any time t, we may draw spheres with radius vt, round every point of the boundary of T These spheres will have one or two bounding envelopes, which separate the space cut by the spheres, from that which includes all points which are not cut by any sphere of radius vt drawn round any point within T as centre. The envelope or envelopes therefore form the boundary of the disturbed region In Figures 19, 20 and 21 the disturbed space is supposed to have a rectangular section, and the sections of those waves are drawn which spread out from the edges of the disturbed region In the first figure the time t is taken to be small, so that there is only one envelope and one boundary In Fig 20, t has increased sufficiently to show a space in the centre of the originally disturbed region, in which equilibrium has been restored. This space spreads out until as shown in Fig. 21 the

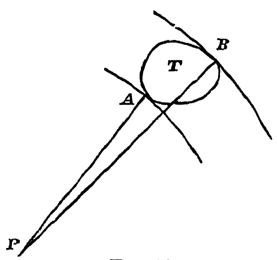

Fig 18

disturbance is confined to a shell, including a considerable space in which the disturbance has ceased, the boundaries of the disturbed region approach the shape of spheres

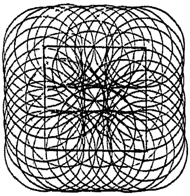

Fig 19

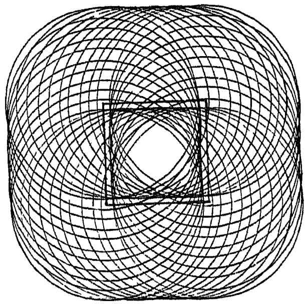

Fig 20

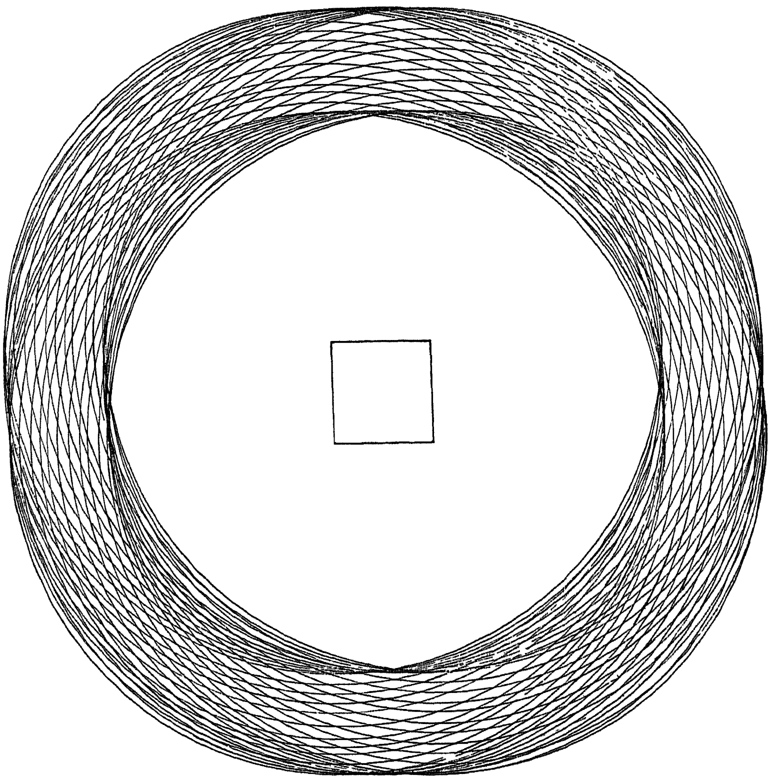

Fig 21

16. The Principle of Superposition. It has been assumed in the last article that the disturbance at P may be obtained by superposing the disturbances reaching it separately from all wave centres within T This is called the principle of superposition, and holds, as may easily be proved, when the elastic properties of the medium are such that the stresses are linear functions of the displacements, or of their differential coefficients with respect to the coordinate axes

In the special cases discussed in Arts 11 and 12, y being the displacement, the stresses are proportional to $\frac{d^2y}{dx^2}$, and satisfy therefore the condition of linearity This is still found to be true if the investigation is not limited to plane waves, for whatever be the properties of the medium, the stresses are always functions of the strains, of such form that when the strains are small, their squares and products may ultimately be neglected The principle of superposition may always therefore be taken to be an approximation which becomes more and more nearly true, the smaller the motion

17 Huygens' Secondary Waves Instead of following a disturbance from its original source, it is often more convenient to trace its subsequent propagation from its position and character at a given time Thus let a disturbance originally coming from a small space be spread at time t over a thin spherical shell of which a portion AB is shown in Fig 22 We may consider this shell to be the disturbed region and find the disturbance at time t_1 from Art 15 by drawing spheres with radius $v(t_1 - t)$ round each point of the shell. We get in this way two spherical envelopes $H'K'$ and HK between which the disturbance is necessarily confined

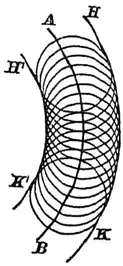

Fig 22

This result seems to be in contradiction with that obtained by another line of reasoning, for, going back to the original cause of the disturbance, the latter should, at time t_1, be confined to a thin shell of which HK is the outer boundary, and except close to HK there should be no disturbance.

This brings us to the important remark that the construction of Art 15 only gives us the space in which there *may* be a disturbance, and not the space in which there is one necessarily When the displacements and velocities of the originally disturbed regions are independent of each other, each point of the space in question will in general have a velocity and a displacement, and only in exceptional

cases will these reduce to zero But the displacements and velocities in the shell AB (Fig. 22) are not independent of each other, for they all originally come from the same source. Hence the waves which we may imagine to spread out from different points of AB must have some relation to each other as regards direction of displacement and velocity. As both our methods of reasoning are correct, it follows that the relation in the present instance must be such that there is neutralization at all points except in a narrow space close to the outer boundary HK.

If we imagine the velocities in AB to be reversed, the displacements remaining the same, we should get a wave travelling inwards In that case, there should be neutralization of the secondary waves over HK and the disturbances would now lie in a shell close up to $H'K'$. This shows that the question whether a wave travels in one direction or another depends on the relation between velocities and displacements The same result has already been proved in Art 11

The propagation of waves not necessarily plane or spherical may be treated in the same manner As long as we know that the disturbance originally comes from a small space, and is therefore confined to a thin sheet, we may always have recourse to the proposition, according to which the disturbance at time t_1 is obtained from that at time t by constructing the outer envelope of all spheres having a radius $v(t_1 - t)$ and their centres on the boundary of the space to which the disturbance is confined at time t

Huygens was the first to investigate the propagation of waves by considering secondary waves to spread out from all points of a disturbance, but the question why the disturbance should be confined to the outer envelope of the secondary spheres has been a serious difficulty up to the time of Fresnel, and even now the reason why, according to Huygens' construction, a wave should not be propagated backwards as well as forwards, is often a stumbling-block

18. Refraction and Reflexion of waves. Imagine a plane wave disturbance to be confined to a narrow layer between two parallel planes of which AB and $A'B'$ are the intersections with the plane of the paper Let this wave meet a surface HK which forms the boundary of another medium having similar properties to the first, but differing in the rate at which the waves travel through it

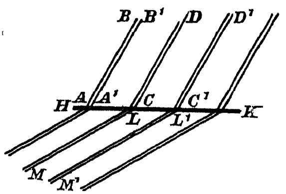

Fig 23

If the second medium had the same velocity of propagation, the waves at subsequent times t_1 would be spread over a space between the parallel sheets CD, $C'D'$, and it will now be shown that the wave on entering the second medium remains a plane wave, but with changed direction, so that LM, $L'M'$ may represent the boundaries of the space to which the wave has spread. To prove this, let AB (Fig. 24) represent the front of the sheet of disturbance which is supposed to be at right angles to the plane of the paper. After a time t the wave has moved forward in the first medium through a distance

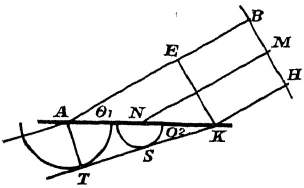

Fig 24

$$BH = v_1 t.$$

In the meantime, we may imagine, according to the previous articles, a secondary wave to have spread from A through a distance $v_2 t$, where v_2 is the velocity of propagation in the second medium. Draw therefore a sphere of radius $AT = v_2 t$. To trace another secondary wave we choose a time, say t/n, at which the wave occupies in the first medium a position such that $BM = BH/n$, its point of intersection with the line AK will be N, such that $AN = AK/n$. From this point N, waves spread out, and at time t, i.e. an interval $t\left(1-\frac{1}{n}\right)$ after the wave has reached N, this secondary wave will have a radius $v_1 t\left(1-\frac{1}{n}\right)$. If all these secondary waves are drawn for values of v_1 between 0 and 1, they are found to have a common tangent plane KST. This tangent plane gives the extreme limit of the disturbance in the second medium at the time t and represents therefore the wave-front at the time t. Draw KE normal to the wave in the first medium, AT normal to the wave in the second medium, and let θ_1 and θ_2 represent the angles between the wave and the surface of separation. Then an inspection of the figure gives at once

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{KE}{AT} = \frac{v_1}{v_2} \quad . \quad . \quad . \quad (9).$$

We call the wave in the second medium the refracted wave, and equation (9) gives the law of refraction. The "refractive index" of a substance as commonly defined is therefore equal to the ratio of the velocity of light in vacuo, to the velocity of light in the substance. The reflexion of waves may be treated exactly in the same manner, and the well-known law deduced, according to which incident and reflected waves are equally inclined to the surface of separation.

19. Wave Front and Wave Surface. In a medium in which waves of all periods are propagated with equal velocities, a wave-front is best defined as a surface such that the disturbance over it originally came from the same source, and started from that source at the same time This does not restrict us to any particular form or shape of the wave If the disturbance follows the law of normal oscillations the wave-fronts are also surfaces of equal phase This follows from the fact that if we imagine ourselves to follow *e g* a condensation as it leaves a source, and spreads outwards with the velocity at which the wave is propagated, the locus of the condensation will, by the above definition, be a wave-front, it will also remain a locus of equal phase and remain so, though the wave may be refracted and reflected When the medium transmits waves of different lengths with different velocities, the above definition no longer applies, but we may still trace the surfaces of equal phase in the case of simply periodic oscillations

A wave-front may lie altogether in one medium, or partly in one and partly in the other Thus in Fig 23 *DCLM* represents the trace of a wave-front We speak of a wave *surface* when we refer to the front of a wave which completely surrounds a small centre of disturbance, and has never passed out of the medium in which that centre lies A wave *surface* in a homogeneous medium like air, glass, or water, is always a sphere, while the shape of a wave *front* would depend on the previous history of the wave, and might be plane, spherical or of irregular shape In all media whether crystalline or isotropic the wave-surface is characteristic of the medium, while the wave-front in general is not

CHAPTER III.

PRELIMINARY DISCUSSION OF THE NATURE OF LIGHT AND ITS PROPAGATION

20 The Nature of Light. We imagine the luminiferous æther to be a medium, filling all space and permeating all bodies Light is a wave-motion in this medium The waves of light are of the nature of distortional waves, the displacements in transparent and isotropic bodies being in the wave-front Waves of the simple periodic form are propagated through the æther with a velocity independent of the wave-length. Hence any plane wave may be propagated without change of type

A wave in which the displacements *at every point* are simply periodic, is called a homogeneous wave If *e g* the displacement in a plane wave travelling in the direction of x is represented by

$$y = a \cos 2\pi \left(\frac{t}{\tau} - \frac{x}{\lambda}\right) \qquad (1),$$

without limitation as to the distance x, we should have a homogeneous vibration of wave-length λ, period τ, and frequency $1/\tau$ But we have no practical experience of a homogeneous wave of light If it existed, *i e* if equation (1) were strictly true, the oscillation of any point would know no limit as regards time, either in the positive or negative direction A particle cannot send out homogeneous radiations unless it has been vibrating for an infinite time, and the mere fact that we are lighting a flame, and extinguishing it, shows that the flame does not send out homogeneous radiations Students should clearly realize that this is a consequence of our definition of homogeneous light We cannot alter that definition without introducing a vagueness into our ideas, which has been the cause of much error and confusion

Our perception of light depends on a physiological sensation, but the waves which are capable of producing this sensation are restricted to a definite range of frequency There are radiations which have all the properties of luminous radiations, but which we cannot perceive by means of our eyes because their wave-length lies outside that range

When we speak of the "spectrum" we include the whole range of radiation emitted by a radiating body, and we distinguish between the visible portion of the spectrum, which extends from the red to the violet, and the invisible portion which includes the wave-lengths which are too long to produce a visible sensation (infrared radiations) and those which are too short to produce a visible effect (ultraviolet radiations) A heated body emits radiations consisting of transverse waves, which when the temperature is low, belong entirely to the infrared portion of the spectrum. As the temperature increases, shorter waves are added to the radiation and increase in intensity both absolutely and relatively to the rays previously emitted Ultimately the waves belonging to the visible portion of the spectrum begin to be included, when the body becomes red hot A still further increase of temperature adds other visible and ultimately the ultraviolet radiations

Table I gives an approximate idea of the length of different waves

TABLE I

	cms
Extreme Infrared radiation observed by Rubens and Aschkinass	0061
Extreme Infrared in Solar Spectrum	00053
,, Red of Visible portion	000077
,, Violet ,, ,,	000039
,, Ultraviolet transmitted through atmospheric air	000018
Extreme Ultraviolet observed by V Schuman	000010

The electrical vibrations emitted by an electric spark are of the same nature as luminous radiations, but the shortest electrical wave we have been able to produce is several millimetres long, i e about one hundred times longer than the longest observed infrared wave

We shall discuss in Chapter XIII the knowledge we possess of the nature of light as it is emitted by incandescent bodies, but for the present it will be sufficient to introduce a simplification which is not in contradiction with any known experimental fact All known sources of light, even those most nearly homogeneous, can be treated as emitting a large number of radiations, each being homogeneous, but differing from each other in wave-length In the case of ordinary white light these vibrations must be distributed throughout the whole range of the spectrum, and must be sufficiently near each other to escape the possibility of resolution by any known spectroscope If the space between wave-lengths $\lambda_1 = 5\,889 \times 10^{-5}$ cms, and $\lambda_2 = 5\,895 \times 10^{-5}$ cms (which are the wave-lengths of the components of the yellow sodium doublet) were filled uniformly by waves, each differing from the other in length by the millionth part of $\lambda_2 - \lambda_1$, no power available, or likely

to be available, would recognise the intervals between these waves, but the light appearing in a spectroscope would seem to fill uniformly the space included between the sodium lines If throughout the spectrum homogeneous vibrations were distributed at intervals equal to the above, no instrument could tell us that we were not dealing with what is called a continuous spectrum. We are at liberty therefore to assume that all continuous spectra are made up of homogeneous vibrations in such close proximity that we cannot separate them This is not put forward as a physical theory, but as a method of obtaining an analytical expression of the facts in a simple manner

21. Intensity. In comparing different radiations in the same medium, we may take the square of the amplitude as a measure of their intensity. As comparative measurements are always made in the same medium, this definition is sufficient for practical purposes Waves of different wave-lengths can only be compared with each other when their energy is converted into some common type This is generally effected by absorption, the heat equivalent of the radiations being compared by the bolometer or thermopile

22 Velocity of Light The experimental methods by means of which the velocity of light may be measured are explained in elementary books Fizeau's method of revolving apertures was used by A Cornu in a series of experiments to which the highest value must be attached The final number arrived at for the velocity in vacuo was

$$3\ 004 \times 10^{10} \text{ cms /second,}$$

a result which is not likely to be in error by more than $3\ {}^0/_0$.

Foucault's method of the revolving mirror was used by Michelson, and later by Newcomb in conjunction with Michelson The final result gave

$$2\ 9986 \times 10^{10} \text{ cms /second}$$

The accidental errors of this method seem considerably smaller than in the method of Fizeau, but certain assumptions on which it rests are not quite free from objection Professor A Cornu has published in the *Rapports de Physique du Congrès International de Physique*, 1900, a very clear discussion of the relative merits of the two methods His conclusion is, that the arithmetic mean of the above determinations gives us at present the best result, and that the most probable value of velocity of light in C G S units is

$$3\ 0013 \times 10^{10}.$$

An error of one part in a thousand in the number is quite possible so that for all purposes we may for convenience adopt the simple and easily remembered number 3×10^{10} The velocity of light in empty space will throughout this book be denoted by V.

23. Optical length and optical distance. The optical length of a path is defined as its equivalent in vacuo, two lengths being called equivalent when light occupies the same time in travelling along them. If the path traverses several media, the total optical length is the sum of the optical lengths of all the different parts. Thus if v_1, v_2, v_3, etc are the velocities of light and s_1, s_2, s_3, etc the lengths of the paths in the various media, then the optical length is

$$V\left(\frac{s_1}{v_1}+\frac{s_2}{v_2}+\frac{s_3}{v_3}+\quad\right)$$

But by Art 18, if μ_1, μ_2, μ_3 are the refractive indices,

$$\mu_1 = V/v_1\,,\quad \mu_2 = V/v_2\,,\quad \mu_3 = V/v_3\,,$$

and hence the optical length of the path is

$$\mu_1 s_1 + \mu_2 s_2 + \mu_3 s_3 + \qquad (2)$$

The optical *distance* between two points is defined to be the shortest optical length of any line, curved, straight, or broken, that can be drawn between them If both points lie in the same medium, the shortest path is clearly the straight line which joins them, and the optical distance is the length of this line multiplied by the refractive index of the substance

A "ray" is defined to be a path of shortest optical length In a medium possessing uniform optical properties, a ray passing through two given points, must, by this definition, always be the straight line which joins them. The path of a ray between two points which are situated in different media may be determined as follows

Let A and R, Fig 25, be the two points, and S some point on the surface of separation, which lies in the plane drawn through A and R, perpendicular to the surface Draw AC perpendicular to AS, and RE perpendicular to SR From any point T in the plane ASR draw TC parallel to AS, and TE parallel to SR, and construct perpendiculars SH and KT from S and T, on CT and SR respectively Let the position of S be such that the optical length HT is equal to the optical length SK, then the optical length of $CT+TE$ is equal to that of $AS+SR$ But from inspection of the figure, $AT>CT$, $RT>TE$, hence the optical length of the path $AT+TR$ must be greater than that of the path $CT+TE$ which is equal to that of $AS+SR$ Students should convince themselves that the same result follows when the point T is taken to lie on the other side of S It follows that the optical length $AS+SR$ is smaller than that of any other path joining A and R in the plane of the paper The condition on which this result depends is that the

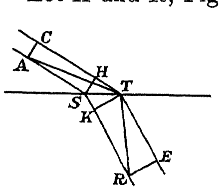

Fig 25

optical length of HT is equal to that of SK or that if μ_1, μ_2 are the two refractive indices,

$$\mu_1 HT = \mu_2 SK$$

If θ_1 and θ_2 are the angles which AS and SR form with the normal to the surface, the condition reduces to

$$\mu_1 \sin \theta_1 = \mu_2 \sin \theta_2,$$

which is the well-known law of refraction The rays as defined by us are therefore identical with the rays of geometrical optics.

It has been assumed in the above proof, that the path of shortest optical distance lies in the plane which is at right angles to the surface separating the two media The restriction may be removed by giving to S a small displacement to either side at right angles to that plane, and showing that the optical distances AS and SR are both clearly increased

A ray may be drawn between any two points of an optical system, but only a single set of rays belong to one set of wave-fronts Let HK and $H'K'$ (Fig 26) represent two wave-fronts of the same disturbance From a point A on HK, a line may be drawn tracing the shortest optical length between A and any given point C on $H'K'$. By altering the position of C, its optical distance from A changes, and some point may be found on $H'K'$ for which that optical distance is least. Let B be that point The path of shortest optical length between A and B is one ray of the system which belongs to the two wave-fronts We may similarly trace a ray satisfying the same conditions from every point P on HK to a corresponding point Q on $H'K'$, and thus obtain the system of rays belonging to a given system of wave-fronts

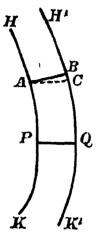

Fig. 26

If the medium is homogeneous, the rays must be straight lines. In a number of separate media, each being homogeneous, the system of rays is made up of a system of straight lines, which will in general change in direction when passing from one medium to another

If the medium is isotropic, so that one wave-front may be obtained from another by Huygens' construction, as explained in Art 17, the system of rays intersects the system of wave-fronts at right angles This is proved by considering two points, A_1, A_2, on a wave-front HK Every other wave-front $H'K'$ will be a tangent surface to two spheres, drawn with the same radius round A_1 and A_2 as centres, so that if B_1, B_2, are the two points of contact, A_1B_1 and A_2B_2 must be at right angles to $H'K'$ This being so, A_2B_1 is necessarily longer than A_1B_1, provided that A_2 is sufficiently near to A_1. Hence all points on HK which are near

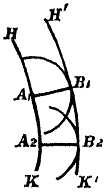

Fig. 27

A_1 are further from B_1 than A_1, and therefore the sphere which is drawn through A_1 round B_1 as centre, cannot intersect, but must touch the surface HK. A_1B_1 stands therefore at right angles both to HK and to $H'K'$.

If the medium is isotropic, but not homogeneous, as *e g* the air surrounding the earth, which varies in density and temperature, the course of a ray may be curved, but the above proof still holds if we take HK, $H'K'$ to lie near each other, and hence the rays are in this case also at right angles to the wave-fronts.

It also follows from Huygens' construction that the optical length from one wave-front to another is the same when measured along different rays. We shall call this length the optical distance between the two wave-fronts.

To illustrate the use which may be made of these propositions, we may deduce the well-known formula connecting the position of a small object with that of its image formed by a lens.

If waves spread out from a point source at P, the wave-fronts are spheres with the point as centre. If these wave-fronts, after passing through the lens, are spheres with Q as centre, the wave-fronts will gradually contract until the energy of the waves is concentrated at Q. (This is not quite correct, owing to the fact that the wave-fronts after emergence are not *complete* spheres, but this does not affect the argument.) The optical length from P to any point on HK is the same, and also the optical length from any point on $H'K'$ to Q. It has been proved above that the optical distance from HK to $H'K'$ is the same when measured along any ray, hence the optical distance from an object to its image is the same along all rays. PSQ and $PMNQ$ are clearly lines satisfying all conditions laid down for the rays belonging to the system. If μ is the refractive index of the lens, the equality of optical lengths leads to the equation

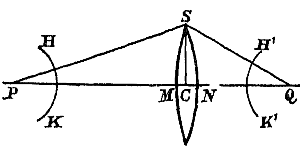

Fig 28

$$PS + SQ = PM + \mu MN + NQ = PQ + (\mu - 1) MN,$$

or

$$(PS - PC) + (SQ - QC) = (\mu - 1) MN.$$

Also

$$PS^2 - PC^2 = SC^2,$$

$$PS - PC = \frac{SC^2}{PS + PC} = \frac{SC^2}{2PC},$$

if the angle SPC is so small that its square may be neglected.

Similarly

$$SQ - QC = \frac{SC^2}{2CQ},$$

$$\frac{1}{PC} + \frac{1}{CQ} = \frac{2(\mu - 1) MN}{SC^2}.$$

If MN and SC are expressed in terms of the radii of curvature of the surfaces of the lens, we obtain the well-known relation between the position of object and image.

24 Fermat's Principle and its application. Fermat (1608 to 1668) making use of the argument that Nature could not be wasteful, and was bound for this reason to cause the rays of light to travel between two points in the shortest time possible, was able to deduce from this proposition the laws of reflexion and refraction. Though we do not now attach any weight to the premiss, we accept the conclusion.

"*Fermat's Principle,*" as it is called, may serve as a connecting link between the waves of the undulatory theory, and the rays of Geometrical Optics, and often gives us a powerful method of dealing quickly with otherwise complicated problems. The ray being *defined* as the path of the shortest optical length, Fermat's principle requires no proof, but what must now be proved, and has been proved above, is, that the course of the rays so defined leads to the correct construction according to laws of geometrical optics. The importance of the property of minimum optical length lies in the fact that it enables us often to determine optical distances with sufficient accuracy when the course of the rays is only approximately known. That the optical length is the same when measured along a ray or a line infinitely near the ray, follows from the minimum property, but in view of the importance of the proposition, it may be more formally proved thus

HK, $H'K'$ being wave-fronts, let APB be a ray belonging to the system. Let AQB be a line lying near APB along its whole course, in such a way that their distance apart ST at any point S may be expressed in terms of the position of S, and the separation PQ at some definite point Q. Writing $PQ = \alpha$, the difference in the optical length of AQB and APB must then be expressible in terms of α, and if α is small, must be capable of expansion in a series proceeding by powers of α as e.g.

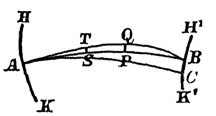

Fig. 29

$$h_1\alpha + h_2\alpha^2 + h_3\alpha^3 +$$

If the product $h_1\alpha$ were negative and α were made to diminish in magnitude, so that the higher powers ultimately vanish, it would follow that AQB is *shorter than* APB, which is contrary to the supposition that APB is a path of minimum length. Hence $h_1\alpha$ cannot be negative, but as nothing limits the sign of α, it follows that h_1 must be zero. We also see that the difference in length between APB and AQB can ultimately only depend on the square of PQ, hence we may conclude

that if a problem deals with the *angles* between adjoining rays, which depend on the first power of PQ, we may neglect when PQ is small any differences in optical length between them. The two adjoining rays need not intersect at the wave-fronts, but they must both start and both end in the same wave-front Thus the difference in length between AC and AB may ultimately be taken to be of the second order of magnitude when the point C approaches B

The application of Fermat's principle may be illustrated by an example

Let a parallel beam (*i e* a beam in which the rays are parallel and therefore the wave-fronts planes which cut the rays at right angles) fall on a prism and be refracted through it Let HK and LM (Fig 30) be two wave-fronts, then μ being the refractive index, the equality of optical lengths gives

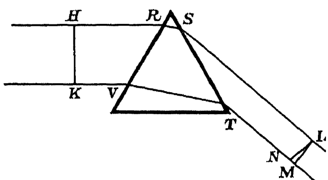

Fig 30

$$HR + \mu RS + SL = KV + \mu VT + TM.$$

Suppose a wave of slightly different wave-length and refractive index μ' falls on the prism, the incident beam being coincident with that just considered We may take HK as being also one of the fronts of the second set of waves, but on emergence, the wave-fronts for the wave defined by μ' will not be parallel to those defined by μ We select that front which passes through L Let its inclination be such that it intersects the ray TM in N If μ' and μ only differ by a small quantity, we may measure the optical length of any of the rays μ' not along its own path, which we do not know, but along the path traced out by one of the rays μ which lies near it, the error committed depending only on the second power of $\mu' - \mu$ We may therefore obtain a second equation for the equality of optical lengths, which is

$$HR + \mu' RS + SL = KV + \mu' VT + TN,$$

taking the difference between the two equations,

$$(\mu' - \mu) RS = (\mu' - \mu) VT - MN,$$

or

$$(\mu' - \mu)(VT - RS) = MN$$

The angle θ formed between the emergent rays of the two beams is equal to the angle NLM, or if small, equal to its tangent NM/ML. It follows that

$$\frac{\theta}{(\mu' - \mu)} = \frac{VT - RS}{ML} \quad . \quad (3).$$

This is a useful expression, first obtained by Lord Rayleigh, connecting the dispersion of a prism with the width of the emergent beam, and the lengths of the paths traversed in the prism by the extreme rays of the beam

25. The Principle of Reversibility. According to an important proposition, a reversal at any time of all velocities in a dynamical system, in which there is no dissipation of energy, leads to a complete reversal of the previous motion Any configuration of the system which existed at a time t before the reversal took place will again exist at the time t after reversal

As an example of this principle, I give an investigation originally due to Stokes, which yields important relations between the amplitudes of incident, reflected and refracted light. Let a ray of homogeneous light AO (Fig 31), of unit amplitude, fall on a reflecting surface Let r be the amplitude of the reflected ray OR, and t that of the transmitted ray OT. If at any moment the courses of the reflected and refracted rays are reversed, the two reversed rays coming together at the surface should combine to reproduce the ray of unit amplitude passing along OA and nothing else That is to say, the ray OT' due to the reflexion of TO must be neutralized by the ray due to the refraction of RO We shall begin by assuming that there is no change of phase at reflexion or refraction, except possibly one of 180° which will appear as a reversal of the sign of the amplitude If r' measures the amplitude after reflexion at O of a ray of unit amplitude travelling along TO, the ray which originally travelled along AO with unit amplitude, and after refraction took an amplitude t, will, after reversal and reflexion at O, have an amplitude tr' Similarly the ray OR reversed and refracted takes an amplitude rt Hence one of the conclusions we may draw from the principle of reversion is that

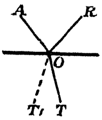

Fig 31

$$rt + r't = 0$$

or

$$r + r' = 0 \tag{4}$$

This equation must be interpreted to mean that there is a reversal of phase either at internal or external reflexion, r' being equal in magnitude to r, but of opposite sign

The ray OR of amplitude r, has after reversal and renewed reflexion at O, an amplitude r^2, the ray OT of amplitude t has, after reversal and refraction at O, an amplitude tt', if t' is the ratio in amplitude of the incident and refracted ray when the ray passes through the surface in the reverse direction If the two rays make up the original one of unit amplitude, it follows that

$$r^2 + tt' = 1 \tag{5}$$

The equations are not sufficient to determine t and t' in terms of r, but they establish an important relation

We may now generalize our results so as to include the possibility of a change of phase.

Let the oscillation in the incident ray at the point of contact with the reflecting surface be given by the projection on a fixed line of the revolution of a point I (Fig 32) in a circle, the revolution being counter-clockwise Let similarly the motion at the same point of the reflected and refracted waves be represented by the projection of the circular motion of R and T The system of points I, R, T revolving with the same angular velocity represents at any time, the phases at the point of incidence of the incident, reflected and refracted rays At some instant let the rays be reversed, the effect on our diagram will be that the points T and R now revolve clockwise, but their position at the time of reversal is unchanged The reflected ray reversed will give a reflected ray represented by OR_1 where R_1 must lie on OI, because obviously the reflexion in the reverse direction must produce a change of phase which is identical in magnitude with the change of phase in the forward direction The refracted ray OT gives rise, on reversal, to a refracted ray, which again must be capable of representation as a projection of the circular motion of some point T_1 and this point must also lie on OI because the principle of reversion shows that OR_1 and OT_1 must have the resultant OI

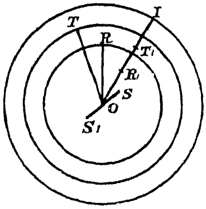

Fig 32

With the same notation as before, we find that the equation

$$r^2 + tt' = 1$$

is independent of any assumption as to change of phase at reflexion

The reflected wave OR gives rise after its reversal to a refracted wave which as regards phase and amplitude, may be represented by OS where $OS = tr$, while the refracted wave OT gives rise to a reflected wave represented by the vector $OS_1 = tr'$, which must neutralize OS. This as regards magnitude leads to the equation

$$r = r'$$

Now the angle ROS must be equal to IOT, and

$$\begin{aligned} TOS &= TOI + IOS \\ &= TOI + ROS - ROI \\ &= 2TOI - ROI \end{aligned}$$

If the change of phase (IOT) at transmission be denoted by τ, and the change of phase at reflexion (IOR) by ρ, then the change of phase ρ' at the internal reflexion is TOS_1, measured clockwise, which is the direction of the reversed motion, this is equal to $\pi + TOS$

Hence $$\rho' = \pi + 2\tau - \rho$$

or $$\rho + \rho' = \pi + 2\tau$$

gives the complete law, which for $\tau = 0$ reduces to $\rho + \rho' = \pi$ as previously established The change of phase at transmission is by the same reasoning shown to be the same in whichever direction the refraction takes place This completes the information we can get out of the principle of reversibility in dealing with this problem

26. Polarization. If in a wave of light, a point of the medium moves in a straight line, we call the light plane polarized If the path of a point is an unchanging ellipse or circle, we speak of elliptical or circular polarization The phenomenon of polarization was first discovered by observing that light could be obtained which showed properties which were unsymmetrical with respect to the ray Thus if a ray of light AR (Fig 33) be reflected from a glass surface HK, at a particular angle depending on the refractive index of the glass, and the reflected ray RS be incident on a second mirror LM, which in one position is parallel to HK, but is capable of rotation round an axis OS coincident with RS, the intensity of the reflected ray SB depends on the position of the second mirror If LM be parallel to HK, the intensity of the reflected ray is a maximum, and if the mirror be turned through a right angle, so that the plane of incidence, instead of being in the plane of the paper, is at right angles to it, the intensity of the reflected light is zero Such a result has no analogy in sound and would not be capable of explanation if light were due to longitudinal waves In the case of transverse disturbances, we may draw a distinction between the vibrations which lie in the plane of incidence and those at right angles to it, and thus explain the want of symmetry If we imagine that at a particular incidence, those rays only are reflected in which the vibration is at right angles to the plane of incidence, the ray RS will consist of vibrations at right angles to this plane and will be reflected in the same proportion by LM, if the two mirrors are parallel But if LM be turned through a right angle, the vibration along RS will now be *in* the plane of incidence of the second mirror, and hence by hypothesis, no light is reflected. Light which has been polarized by reflexion, is said to be polarized *in* the plane of incidence

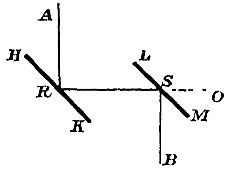

Fig 33

All homogeneous rays are polarized To prove this, we imagine the wave to proceed in the direction of the axis of x.

Let the displacement have one component along OY,

$$y = a \cos \omega t,$$

and one along OZ,

$$z = b \cos \omega (t - \theta)$$

According to Art. 6 the motion is rectilinear when $\theta = 0$, circular whe $a = b$ and $\theta = \pm \pi/2$, and elliptical in all other cases Homogeneou light may therefore be plane, circularly, or elliptically polarized, but i will always be polarized.

We have certain experimental methods of detecting polarizatio When these methods are applied to light emitted from a flame or fro a body rendered incandescent by the electric discharge, it is found tha under ordinary circumstances no polarization can be detected, eve when a single spectral line is examined We conclude that the ligh emitted from these sources is not rigorously homogeneous, though it i often called so The range of wave-length may be small in thes cases, but it is not infinitely small.

If we project a spectrum on a screen, and by means of a narro slit in the screen, separate a small portion of the spectrum, we shal find that this portion is not polarized We must conclude that th vibrations are not strictly homogeneous, though we may have separate from the spectrum waves the extremes of which differ only very littl in period Let λ_1, λ_2, λ_3 λ_n, be wave-lengths in such clos proximity to each other, that even the most powerful availabl spectroscope cannot separate λ_1 from λ_2 Let each wave be homo geneous and therefore polarized Assuming plane polarization for th sake of simplicity, there need be no relationship between the direction of vibrations of the different periods We may define the direction o vibration by the angle α between it and a fixed direction in the plan of the wave If the vibrations are irregularly distributed, the numbe lying between directions α and $\alpha + d\alpha$ will be independent of α Henc in the resulting motion formed by the overlapping of all the trains o waves considered, there can be no preference for any one particula direction, and the light will appear to be unpolarized This want o polarization is not an intrinsic property of the light emitted by th source, but a consequence of the want of homogeneity, it is caused b the superposition of a large number of polarized oscillations of slightl different periods The difference in the period of the superpose vibrations, however small, is a necessary condition of the want o polarization, because any number of oscillations of the *same* perio would necessarily combine into a motion which must be eithe rectilinear, circular, or elliptical

We may observe polarization in light which is not homogeneous, i the motion for all frequencies is along similar paths, similarly situated or at any rate, if the chance of a number of elliptical vibrations havin its major axes in directions lying between α and $\alpha + d\alpha$ is greater fo some values of α than for others In the former case, we have com plete, and in the latter, partial polarization

27 Light reflected from transparent substances It will be useful to follow out a little more closely at this stage the effects of reflexion from a transparent polished surface According to the preceding article, ordinary light reflected by such a surface at a particular angle, called the angle of polarization, is plane polarized and by definition, polarized in the plane of incidence. This is true whether the incident beam is polarized or not. I anticipate the results of later Chapters by specifying at once, that the direction of vibration is at right angles to what has been called the plane of polarization

The amplitude of the reflected light must, according to what has been said, depend (1) on the direction of polarization of the incident light, and (2) on the angle of incidence. A mathematical expression for the reflected amplitudes in different cases was first obtained by Fresnel, whose results we here adopt, deferring to a later stage a discussion of their justification If a homogeneous vibration of unit amplitude vibrating normally to the plane of incidence falls on a reflecting transparent substance, the angle of incidence being θ, the amplitude of the reflected ray is

$$r_n = \frac{\sin(\theta_1 - \theta)}{\sin(\theta_1 + \theta)} \qquad (6)$$

If the light vibrates parallel to the plane of incidence the reflected vibration has an amplitude

$$r_p = \frac{\tan(\theta_1 - \theta)}{\tan(\theta_1 + \theta)} \qquad (7)$$

In these equations θ_1 denotes the angle of refraction so that if μ is the refractive index, $\sin\theta = \mu \sin\theta_1$ For the present we take these equations to represent experimental facts and apply them to study the polarization effects in particular cases

The square of r_n increases with increasing incidence from $\theta = 0$ (normal incidence) to $\theta = \pi/2$ (grazing incidence). When θ is sufficiently small, we may put $\sin\theta = \theta$, $\theta = \mu\theta_1$, and obtain

$$r_n = \frac{1-\mu}{1+\mu} \qquad (8)$$

This holds for normal incidence and gives us the intensity of the reflected light at that incidence:

$$\left(\frac{\mu - 1}{\mu + 1}\right)^2.$$

Thus for glass with refractive index 1 5, one-ninth part is reflected at normal incidence, and hence eight-ninths are transmitted When the incident ray is as oblique as possible, the light is entirely reflected, none being transmitted The negative sign of r_n when μ is greater than one indicates a change of phase of 180° The expression for the

light polarized at right angles to the plane of incidence, diminishes from

$$r_p = \frac{1-\mu}{1+\mu},$$

for normal incidence, to 0, when $\theta + \theta_1 = \pi/2$ In that case, $\sin\theta_1 = \cos\theta$ and the equation of refraction $\sin\theta = \mu\sin\theta_1$ becomes $\tan\theta = \mu$ If the angle of incidence further increases, the amplitude increases again and for grazing incidence the light is in this case also totally reflected.

Equations (6) and (7) preserve their numerical value, but reverse their sign when θ and θ_1 are interchanged This shows that on reversal of the ray the same fraction of light is reflected, but that if in one case there is no change of phase, a change of 180° takes place in the other case This agrees with the result independently deduced in Art 25

If the incident light has an amplitude a and is polarized in a plane inclined at an angle α to the plane of incidence, we may decompose the oscillations into two, one $a\cos\alpha$ being polarized in the plane of incidence and the other $a\sin\alpha$ polarized at right angles to that plane The reflected rays of each component may then be united again If b be the amplitude of the reflected ray, and β the angle its plane of polarization forms with the plane of incidence, we have

$$b\cos\beta = a\cos\alpha\frac{\sin(\theta_1-\theta)}{\sin(\theta_1+\theta)},$$

$$b\sin\beta = a\sin\alpha\frac{\tan(\theta_1-\theta)}{\tan(\theta_1+\theta)}$$

Hence
$$\tan\beta = \tan\alpha\frac{\cos(\theta_1+\theta)}{\cos(\theta_1-\theta)},$$

and
$$b^2 = a^2\left\{\cos^2\alpha\frac{\sin^2(\theta_1-\theta)}{\sin^2(\theta_1+\theta)} + \sin^2\alpha\frac{\tan^2(\theta_1-\theta)}{\tan^2(\theta_1+\theta)}\right\}.$$

The first of these equations shows that for $\theta + \theta_1 = \pi/2$, the reflected ray is polarized entirely in the plane of incidence The resulting value of θ obtained from $\tan\theta = \mu$, gives us therefore the angle of polarization

When ordinary light falls on a reflecting surface, we may obtain the intensity of the reflected light by considering that the homogeneous waves of closely adjoining wave-lengths have their planes of polarization distributed quite irregularly, $\cos^2\alpha$ and $\sin^2\alpha$ in the above equation must therefore be replaced by their average value, which is one half The intensity of the reflected light is therefore

$$b^2 = \tfrac{1}{2}a^2\frac{\sin^2(\theta_1-\theta)}{\sin^2(\theta_1+\theta)}\left(1 + \frac{\cos^2(\theta_1+\theta)}{\cos^2(\theta_1-\theta)}\right).$$

The total intensity is given by this expression, but the intensity is distributed unsymmetrically in different directions That part of the

light which is polarized at right angles to the plane of polarization has an intensity

$$\tfrac{1}{2}a^2 \frac{\tan^2(\theta_1 - \theta)}{\tan^2(\theta_1 + \theta)},$$

while for the light polarized in the plane of incidence, the intensity is

$$\tfrac{1}{2}a^2 \frac{\sin^2(\theta_1 - \theta)}{\sin^2(\theta_1 + \theta)}$$

The difference between these two quantities gives us the amount of polarized light, which, together with the unpolarized light of intensity equal to twice the smaller, makes up the partially polarized beam of the reflected light

The intensities of the transmitted beams are obtained by the principle of the conservation of energy, and if I_a, I_r, I_t, represent the intensities of the incident, reflected, and transmitted beams respectively,

$$I_a = I_r + I_t$$

It would be wrong to conclude from this that if a_a, a_r, a_t, measure the *amplitudes* of the incident, reflected, and transmitted rays, $a_a^2 = a_r^2 + a_t^2$, because the squares of amplitudes only express the relative intensities if the waves have the same wave-length, and are transmitted through media possessing the same inertia It is, however, in every case the intensity that concerns us, and the equations given above give therefore everything that is required

As far as can be judged by experiment, Fresnel's equations (6) and (7) very approximately represent the observed facts The most important case is that in which the angle of incidence lies near the angle of polarization If the incident light be polarized at right angles to the plane of incidence, and falls on the surface at the angle of polarization, no light should according to equation (7) be reflected at all, and there should be a complete reversal of phase in the reflected light, as the angle of incidence changes from a value slightly smaller than the angle of polarization to a value slightly greater Sir George Airy discovered that this is not quite correct for highly refracting substances like diamond, and Jamin, pursuing the subject further, found that there is always a residue of light reflected at the polarizing angle though the incident light may be strictly polarized at right angles to the plane of incidence The phase, which should change suddenly through 180°, changes rapidly but not discontinuously, so that at the polarizing angle there is a retardation or acceleration of phase amounting to 90°

Since then, Lord Rayleigh* has shown that Jamin's results are in great part, though not entirely, due to surface films of probably greasy matter which may be removed by polishing.

* *Collected Works*, Vol II p 522.

If light falls on the surface of a plate of glass at the polarizing angle, the ray entering the glass falls on the second surface again at the polarizing angle, as the condition $\theta + \theta_1 = \pi/2$ will, in a plate bounded by parallel surfaces, be fulfilled at both incidences It follows that the light, reflected at the second surface, increases the intensity without detracting from the polarization of the reflected beam The same argument may be used to show that a pile of parallel plates gives at the proper angle a polarized reflected beam which, neglecting absorption, might be made to equal the intensity of that component of the incident beam which is polarized in the plane of incidence Such a pile furnishes a simple and cheap method of obtaining polarized light. There is some disadvantage, however, in the fact that the direction of the rays is changed by reflexion For this reason, the transmitted beam is occasionally used The transmitted beam is only partially polarized by a single refraction, but it is clear that when the number of plates is sufficiently great to reflect all the light polarized in the plane of incidence, the refracted beam can only contain light polarized at right angles to that plane A large number of plates is however required, if the polarization is to be approximately complete The amount of light transmitted through a pile of plates, or reflected from it, has been calculated by Provostaye and Desains*

If ρ be the fraction of the intensity reflected at one surface, that reflected from a number n of parallel surfaces is

$$\frac{n\rho}{1+(n-1)\rho}$$

If there are m plates, there are $2m$ surfaces, hence in terms of m, the intensity of the reflected light is

$$\frac{2m\rho}{1+(2m-1)\rho}$$

and the intensity of the transmitted light is

$$\frac{1-\rho}{1+(2m-1)\rho}$$

For glass of refractive index 1 54, ρ at the polarizing angle is 16 and from this we may calculate that it requires 24 plates to furnish a transmitted beam which shall contain not more than 10 % of unpolarized light

28 Total reflexion. When a ray is totally reflected, there is no refracted ray, but equations (6) and (7) still hold, provided we give to the angle of refraction the imaginary value which it takes according to the laws of refraction, interpreting amplitude, whe

* *Ann de Chemie et Phys* xxx. p. 159 (1850)

it contains an imaginary term, according to principles explained in Art 8 If θ denote the angle of incidence, in a medium of refractive index μ, the second medium being air, the law of refraction is

$$\mu \sin \theta = \sin \theta_1$$

and total reflexion takes place if $\sin \theta > 1/\mu$ In that case we may separate the imaginary and real parts for light vibrating normally to the plane of incidence as follows :

$$\sin (\theta_1 - \theta) = \sin \theta_1 \cos \theta - \cos \theta_1 \sin \theta,$$

$$\sin (\theta_1 + \theta) = \sin \theta_1 \cos \theta + \cos \theta_1 \sin \theta\,,$$

$$r_n = \frac{\sin \theta_1 \cos \theta - \cos \theta_1 \sin \theta}{\sin \theta_1 \cos \theta + \cos \theta_1 \sin \theta}$$

All quantities are real except $\cos \theta$ The expression for r_n is of the form $(p - iq)/(p + iq)$ and hence, according to Art 8, the amplitude is one This was to be expected, since we are dealing with total reflexion

Under these conditions, the complex amplitude is of the form $e^{i\delta_1}$, and its real part measures the cosine of the change of phase (δ_1) The real part of $(p - iq)/(p + iq)$ being $(p^2 - q^2)/(p^2 + q^2)$ we find, with the help of (9)

$$\cos \delta_1 = \frac{1 + \mu^2 - 2\mu^2 \sin^2 \theta}{\mu^2 - 1} \qquad (9).$$

As special cases we have

$$\text{for } \sin \theta = \frac{1}{\mu}, \quad \delta_1 = 0,$$

$$\text{for } \sin \theta = \frac{\pi}{2}, \quad \delta_1 = \pi$$

This shows that at incipient total reflexion there is no change of phase and at grazing incidence, a reversal of phase

In order to reduce the tangent formula, we transform as follows

$$\frac{\tan (\theta_1 - \theta)}{\tan (\theta_1 + \theta)} = \frac{\sin 2\theta_1 - \sin 2\theta}{\sin 2\theta_1 + \sin 2\theta}$$

$$= \frac{(\sin 2\theta_1 - \sin 2\theta)^2}{\sin^2 2\theta_1 - \sin^2 2\theta}$$

Here $\sin 2\theta_1$ is imaginary, but its square is real, hence for the real portion of the fraction we have

$$\cos \delta_2 = \frac{\sin^2 2\theta_1 + \sin^2 2\theta}{\sin^2 2\theta_1 - \sin^2 2\theta}$$

$$= \frac{\mu^2 \cos^2 \theta_1 + \cos^2 \theta}{\mu^2 \cos^2 \theta_1 - \cos^2 \theta},$$

or finally

$$\cos \delta_2 = \frac{(\mu^2 + 1) - (\mu^4 + 1) \sin^2 \theta}{(\mu^2 - 1) - (\mu^4 - 1) \sin^2 \theta} \qquad (10).$$

As special cases we have

$$\text{for } \sin\theta = \frac{1}{\mu}, \quad \delta_2 = \pi,$$

$$\text{for } \sin\theta = \frac{\pi}{2}, \quad \delta_2 = 0$$

The difference in phase of the two components is best obtained directly by taking the real part of r_n/r_p which is equal to $\cos(\delta_1 - \delta_2)$.

But $$\frac{r_n}{r_p} = \frac{\sin(\theta_1 - \theta)}{\sin(\theta_1 + \theta)} \frac{\tan(\theta_1 + \theta)}{\tan(\theta_1 - \theta)} = \frac{\cos(\theta_1 - \theta)}{\cos(\theta_1 + \theta)}$$

$$= \frac{\cos\theta\cos\theta_1 + \sin\theta\sin\theta_1}{\cos\theta\cos\theta_1 - \sin\theta\sin\theta_1}$$

As the only imaginary quantity is $\cos\theta_1$, the expression is of the form

$$\frac{p + iq}{-p + iq}$$

the real part of which is $(q^2 - p^2)/(q^2 + p^2)$

Hence

$$\cos(\delta_1 - \delta_2) = \frac{\cos^2\theta\cos^2\theta_1 + \sin^2\theta\sin^2\theta_1}{\cos^2\theta\cos^2\theta_1 - \sin^2\theta\sin^2\theta_1}$$

$$= \frac{1 + 2\mu^2\sin^4\theta - (1 + \mu^2)\sin^2\theta}{1 - (1 + \mu^2)\sin^2\theta} \qquad (11)$$

Equations (9) and (10), though giving us the values of the change of phase, are unable to distinguish between an acceleration or retarda tion, and equation (11) does not tell us which of the two vibrations i ahead of the other This ambiguity cannot be solved by the mer transformation of Fresnel's formulae We may, it is true, show b means of (11) that $\delta_1 - \delta_2$ does not pass through zero, and hence reaso that if δ_1 is positive, δ_2 must be negative, but recourse must be had t the complete dynamical theory in order to decide which component i accelerated Though the subject has often been treated by variou writers, it was only quite recently that Lord Kelvin* pointed out fo the first time that the vibration in the plane of incidence is retarde while the normal vibration is accelerated, and also that the differenc of phase with the materials at our disposal, is always an obtuse angl The latter conclusion may be derived from equation (11) as it is readil shown that the numerator within the range of total reflexion is positiv and the denominator negative

At incipient total reflexion (where $\mu\sin\theta = 1$) and for grazin incidence, there is a phase difference of 180° Between these tw limits of θ there is one angle for which the difference in phase least This angle is obtained from (11) by putting the differenti

* *Baltimore Lectures*, p 400

coefficient of the right-hand side with respect to $\sin^2\theta$ equal to zero. This gives

$$(\mu^2+1)\sin^2\theta = 2\,,$$

the corresponding maximum retardation is

$$\cos(\delta_1-\delta_2) = -\frac{6\mu^2-\mu^4-1}{(\mu^2+1)^2} \quad \ldots \quad (12)$$

An important practical application of these results was made by Fresnel If it were possible to make the right-hand side of (11) equal to 0, there would be a phase difference of a right angle, which, if the original light was polarized at an angle of 45°, so as to make both components equal, would give circularly polarized light (Art 6) Among the media at our disposal, there is none with a refractive index sufficiently high to give a difference of phase as small as $\pi/2$, but we can secure circularly polarized light by means of two successive reflexions, if $\delta_1-\delta_2=3\pi/4$

Equation (11) may then be written

$$2\mu^2\sin^4\theta-(1+\sqrt{\tfrac{1}{2}})\,\{(1+\mu^2)\sin^2\theta-1\}=0.$$

This is a quadratic equation which may be solved, and has in general two roots Thus for glass of refractive index 1 5, 1 55 and 1 6, the following table gives the calculated values of the two solutions θ_1 and θ_2

TABLE II

$\mu =$	1 5	1 55	1 6
$\delta_1-\delta_2=135°$	50° 14′ 53° 13′ 5	45° 14′ 5 57° 5′	44° 21′ 56° 41′

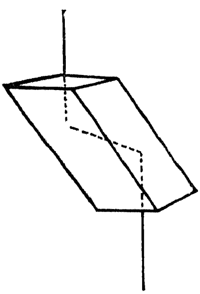

Fig 34

Fresnel's rhomb is a rhomb of glass (Fig. 34) which gives circularly polarized light after two total reflexions in the manner described

Of the two possible angles for the rhomb, the larger is chosen, because it gives a smaller error for slight changes in refrangibility or deviations from the theoretically correct incidence

CHAPTER IV.

THE INTERFERENCE OF LIGHT.

29. The Interference of Light. Under the name of interference of light, we group together all phenomena in which two or more rays, coming originally from the same source, are brought together so as to cause a combined disturbance This resulting disturbance may be calculated if we accept the principle of superposition which has been explained in Art 15 According to that principle, the displacements or velocities produced by any number of centres of disturbance, are obtained by superposing the displacements or velocities due to each. This fundamental principle in Optics rests on a very strong experimental foundation Already Huygens saw the importance of the fact that the passage of a beam of light through an aperture is in no way affected by the passage of another beam through the same aperture As he pointed out, different people may look at different objects through the same opening without noticing any blurring due to the overlapping of the large number of waves which must pass through the opening The waves cross each other at the aperture without in the least interfering with each other's course

The foundation of the principle of superposition rests therefore on the experimental fact, that there is *no* interference between different waves of light, and yet certain phenomena are observed and explained by this principle of non-interference, where the combined effect of two waves differs from the sum of the separate effects The apparent discrepancy arises from the fact that we do not observe the displacements or velocities, but the average squares of the velocities, and the principle of superposition does not hold for the squares of either the displacement or the velocity

There may be therefore, interference of intensity, though there is no interference of displacement To make this clear it is only necessary to refer to Art 4, where it has been shown that the superposition of two periodic motions of the same frequency, amplitudes a_1, a_2, and phase difference δ, leads to a periodic motion, having for its amplitude the square root of

$$a_1^2 + a_2^2 + 2a_1a_2 \cos \delta$$

The last term is a measure of the "interference" as regards the square of amplitude If $a_1 = a_2 = a$, the amplitude becomes

$$2a \cos \frac{\delta}{2}$$

If $\delta = 0$, the amplitude is doubled, but the intensity, which depends on the square of the amplitude, is increased fourfold If δ is equal to two right angles, the amplitude is zero, which means that two oscillations of equal amplitude may neutralize each other After these preliminary remarks, we may show how an experimental illustration may be obtained of this interference of intensities, which it must be always remembered is a consequence of the non-interference of displacements

30 Calculation of the combined effects due to two separate sources. Let P and Q (Fig 35) represent two particles which are sending out waves, the motion at P and Q being simply periodic Let the vibrations be in a direction perpendicular to the plane of the diagram and identical as regards amplitude and period, and let the phases at P and Q be the same Consider a point S on a distant screen, the plane of which is parallel to the line PQ and perpendicular to the plane of the diagram

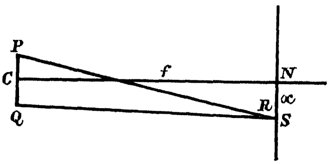

Fig 35

The two motions produced by P and Q at S, considered as acting separately, are parallel to one another, since they are both perpendicular to the plane of the diagram, and they have also approximately the same intensity, if the distance of the screen from the two sources is great compared with the distance of the two sources from one another There will be a difference of phase between the two vibrations due to difference of the distances PS and QS If R be a point on PS, such that $PR = QS$, the phase at R of the vibration transmitted along PS, must be the same as the phase at S due to the vibration transmitted along QS Hence the difference of phase between the two vibrations at S, will be $\frac{2\pi}{\lambda}(PS - QS)$.

Let C be the middle point of PQ, and from C draw CN perpendicular to the plane of the screen, and cutting it in the point N

Let $\qquad CN = f, \ NS = x$, and $PQ = c$

Then
$$PS^2 = f^2 + \left(x + \frac{c}{2}\right)^2 \qquad \ldots \ (1),$$

$$QS^2 = f^2 + \left(x - \frac{c}{2}\right)^2 \qquad \ldots \ (2)$$

Hence $$PS^2 - QS^2 = 2cx.$$

Therefore $$PS - QS = \frac{2cx}{PS + QS}.$$

If x is small compared to f, we may write $2f$ instead of $PS + QS$, the error committed being of the order of magnitude x^2/f^2.

The difference of phase between the two vibrations at S is therefore

$$\frac{2\pi}{\lambda} \frac{xc}{f}$$

Let a denote the amplitude which would be produced at S by each source acting separately

Then the resultant amplitude at S, due to both sources, is

$$2a \cos\left(\frac{\pi}{\lambda} \frac{xc}{f}\right) \qquad (3).$$

The amplitude is variable and depends on the angle x/f

Thus considering points situated on the line NS on the screen, the point N for which $x = 0$, is a point of maximum intensity

The intensity at points on either side of N diminishes symmetrically, and becomes zero when $x = \pm \frac{\lambda f}{2c}$ After this the intensity increases and reaches a maximum again when $x = \pm 2 \frac{\lambda f}{2c}$

The points of maximum intensity are at equal distances $\frac{\lambda f}{c}$ apart and the points of minimum or zero intensity lie halfway between the points of maximum illumination

So far only those points have been considered which lie in the plane PQN, but there is no difficulty in including points outside the central plane. If a point T be taken vertically over S, and at a distance z from it,

$$PT^2 = PS^2 + z^2$$

$$QT^2 = QS^2 + z^2$$

$$\therefore \quad PT^2 - QT^2 = PS^2 - QS^2$$

As long as PT is, to the approximation required, equal to PS, *i e.* as long as z^2 is neglected,

$$PT - QT = PS - QS$$

Hence the illumination at T is the same as the illumination at S, and the illumination of a screen placed at NS consists therefore of a system of alternately bright and dark rectilinear bands, which are at right angles to the plane PQN

If the distance of the screen is altered, the distance of the bands diminishes in direct proportion to the distance from the source, and all bands for which the difference in optical length $PS - QS$ is the same, lie in a plane at right angles to the plane of the paper, and passing through CS

These results require some qualification as they depend on the squares of x and z being neglected The complete investigation is not, however, difficult The locus of the surfaces of equal difference in phase is determined by the condition that $PS - QS$ is a constant, a condition which defines the surfaces as hyperboloids of two sheets having P and Q as foci The intersections of these hyperboloids with the plane of the screen are hyperbolas, and not straight lines, as found by the approximate method, but when the distance between P and Q is small, and those bands only are considered which are situated near the centre of the screen, the hyperbolas are very slightly curved, and may, near the central plane, be considered to be straight lines

If the two sources of light are not in the same phase, but vibrate with a difference of phase which remains constant, the interference bands are formed as before, but the whole system is shifted to one side

Let the vibration emitted from the source P be represented by $a\cos\left(2\pi\frac{t}{T}+\alpha\right)$ and that from Q by $a\cos 2\pi\frac{t}{T}$ Then at a point S on the screen, the difference of phase, which before was $\frac{2\pi(PS-QS)}{\lambda}$ will now be $\frac{2\pi(PS-QS)}{\lambda}+\alpha$ and the position of the bands will be given by the equation

$$PS - QS = \frac{n\lambda}{2} - \frac{\alpha\lambda}{2\pi}$$

The bands are still at the same distance $\lambda f/c$ apart, but the whole system is displaced sideways by an amount equal to $\lambda f\alpha/2\pi c$ The assumption that the oscillations at P and Q are perpendicular to the plane PQN, may be removed, provided that these oscillations are parallel to each other, for under experimental conditions the distance PQ is so small compared with the distance of the screen, that the inclination between the displacements at S caused by parallel disturbances at P and Q may be neglected.

31 Conditions necessary for the experimental illustration of interference. Two homogeneous sources radiating from two points near each other, would, according to the last paragraph, produce a pattern of unequal illumination on a screen, but the position of the bands of maximum and minimum illumination could not *à priori* be

determined unless we knew the differences in phase between the oscillations of the sources. It has been explained in Art 20, that all available sources of light may be supposed to give a number of homogeneous radiations agreeing closely, though not completely, in frequency. The difference in phase between each pair of radiations having equal frequency is quite arbitrary, so that the interference patterns for each of the closely adjacent wave-lengths are quite independent of each other, the dark bands of one wave-length overlap the bright bands of another wave-length, and the illumination of any point of the screen, being the average of a number of superposed effects, is uniform.

To produce visible interference effects, it is necessary that the phase differences between the oscillations of adjacent frequencies should be nearly identical. This cannot be secured if the sources of light are independent, and hence such *independent sources cannot be made to produce interference effects.*

The experimental conditions of interference are obtained by deriving the oscillations originally from the same source.

32. Young's experiment. Both on account of its historical importance and the simplicity of its arrangement, Young's experiment deserves the first place. Two small apertures P and Q (Fig 36) were illuminated by light which originally had passed through another aperture at O. After passing through P and Q, the waves spread out in all directions, and falling on the screen SS', produce equally spaced interference bands. If P and Q are equidistant from O, the phase at P and Q will be the same, hence the central band will be at N. The equality of phase at P and Q holds for waves of all frequencies, and therefore the experimental conditions of Art 31 are realized, and (3) correctly represents the distribution of amplitudes.

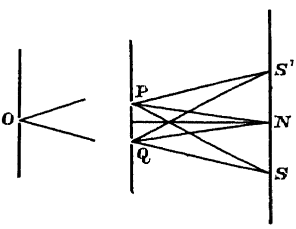

Fig 36

As the distance between the bands depends on the wave-lengths, the light should be nearly homogeneous, if it is desired to observe the effects under the simplest conditions. A great number of bands may thus be seen. To give an idea of the scale on which the experiment has to be conducted, we may take as an example, the distance between P and Q to be 1 mm and the distance of the screen from the aperture to be 1 metre. The distance between the bands is then for red light

$$\frac{\lambda f}{c} = \frac{6 \times 10^{-5} \times 100}{1} = 06 \text{ cms}$$

and similarly for blue light 04 cms The bands are therefore very close together, if we wish to space them further apart, either the distance of the screen has to be increased, or the apertures have to be put closer together

33. Fresnel's experiments In Fresnel's celebrated experiments, the two dependent sources were secured by forming two vertical images of a narrow illuminated slit

Fresnel's Mirrors In the first of the two methods to be described, two inclined mirrors were used to obtain the vertical images.

In Fig 37 OM_1 and OM_2 represent two plane mirrors, which have their planes at right angles to the plane of the diagram Two images A and B are formed by reflexion of the light coming from S

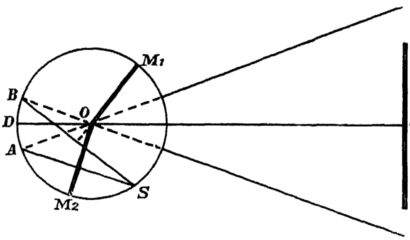

Fig 37

The distance between the two images depends upon the angle of inclination of the two mirrors Let D be the middle point of AB and let DO be produced to meet a distant screen in C Then C will be the centre of the system of interference bands, formed upon the screen To calculate the distance between the two images A and B, we note that A being the image of S formed by the plane mirror OM_2, the distance of A to any point on OM_2 is the same as the distance from S to the same point Hence $OS = OA$ Similarly $OS = OB$ Hence the points A, B and S lie upon a circle with centre at O

Hence $$\angle AOB = 2\angle ASB$$
$$= 2\omega,$$

where ω is the angle between the two mirrors

Therefore $$\angle BOD = \omega$$

Now let $$OS = b \text{ and } OC = d$$

Then $$DO = b\cos\omega \text{ and } DC = d + b\cos\omega$$

Also $$AB = 2BD = 2b\sin\omega$$

The distance between the bands produced on the screen by two sources of light, has been proved to be $\lambda f/c$

In this case $$f = DC = a + b\cos\omega,$$
and $$c = AB = 2b\sin\omega$$

Therefore the distance between the bands produced by Fresnel's mirrors is $\dfrac{\lambda(a + b\cos\omega)}{2b\sin\omega}$ or, since ω is a small angle, $\dfrac{\lambda(a+b)}{2b\omega}$

Fresnel's Biprism In the second method, the two images are obtained by doubling a single source by means of refraction Suppose two similar small-angled prisms OPR, OQR are placed base to base as in the figure This constitutes what is termed Fresnel's Biprism If a source S is placed symmetrically behind the two prisms, two virtual images of it are formed say at A and B

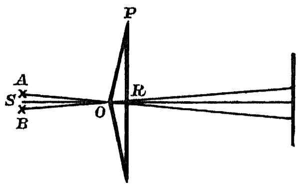

Fig 38

To calculate the distance between the bands, we make use of the fact that a prism of small angle α deviates any ray which falls on one of the faces in a direction nearly normal to it by a quantity $(\mu - 1)\alpha$, where α is the angle of the prism Hence the vertical images of the slit are at the same distance from the prism as the object, and if b be the distance of the slit from the prism, $2(\mu - 1)b\alpha$ measures the distance between the vertical images If a be the distance of the screen from the slit, the general expression for the distance between the bands reduces to

$$\frac{\lambda a}{2b(\mu - 1)\alpha}$$

It should be noticed that the distance between the vertical images in this case, which represents the distance between the two sources of light producing interference, depends on the refractive index, and therefore on the wave-length Plate II Fig 1 is a photograph of the interference bands formed by Fresnel's biprism The rhythmic variation in the intensity of the bands is due to a diffraction effect which will be further alluded to in Art 36

34. Subjective method of observing interference bands. When interference phenomena are observed on a screen in the manner described, the bands are very close together, unless the screen is at a considerable distance from the sources, and in that case, a strong light has to be used if the bands are to be seen There is, however, no difficulty in magnifying the bands by optical means It has been shown in Art 22 that the optical distance between object and image formed by a lens, is the same when measured along all rays If therefore the screen be removed, and the rays crossing at any point P be

focussed by a lens on another screen, the difference in phase between the two rays at the geometrical image of P is the same as the difference in phase at P The interference pattern on the second screen will therefore be an image of the interference pattern on the original screen If the lens in this argument is represented by the focussing arrangement of the eye, so that the retina represents the second screen, the interference effects will be seen just as if they were projected on a screen coincident with the plane for which the eye is adjusted We can also interpose between the eye and the plane for which the eye is focussed a magnifying glass or eye-piece, and this enables us to measure the distance between the bands, for we may introduce a movable cross-wire in the focal plane of the eye-piece This is practically Fresnel's arrangement, and the one which is generally adopted now

If the two slits of Young's arrangement are used, and a telescope focussed for infinity placed close to them, the interference pattern at the focus of the telescope is the image of that which would be formed at infinity, were the telescope away We may use therefore such a telescope to observe the bands

The simplest mode of seeing Young's interference bands has been described by Lord Rayleigh* Two plates of glass are silvered, a fine line is ruled on one of them, and two fine parallel lines, as close together as possible, on the other The ruling of the lines removes the silver film, so that we have now two opaque plates, one containing one slit, and the other, two slits close together. If the double slit is placed close up to the eye, and the other a short distance from it, interference bands are seen when the two plates are so adjusted that their slits are nearly parallel The whole arrangement is easily constructed, and can be mounted in a tube

35 Observations with white light When a nearly homogeneous source of light is used, such as a sodium flame, a large number of bands is easily observed, but with white light only a few bands can be seen, as they soon become indistinct and resolve themselves into a general illumination The reason of this is easily seen from the formula given The distance between the bands depends on the wave-length Under ordinary circumstances the wave-lengths which make an impression on the eye, cover such a range that the length of the extreme violet is about two-thirds of that of the extreme red

Equation (3) shows that for $x=0$ there is a maximum of light for all wave-lengths, the central band appears therefore white As x is increased, a point will be reached at which $x=\lambda f/c$ is satisfied if λ represents the wave-length of the shortest visible, *i e* the violet,

* *British Association Report*, 1893, and *Collected Works*, Vol IV. p 76.

rays. Call the value of x at that point, x_0 For another value, $x=\frac{3}{2}x_0$, the maximum for the red waves will be reached, and for intermediate values we have maxima of intermediate colours On either side of the central band there are therefore a number of coloured spectra The visible colour begins near the point at which the blue has its first minimum and is therefore of a reddish hue

The second maximum of the violet will be reached when $x=2x_0$, the second maximum of the red, when $x=3x_0$, but this same value of x which gives us the second maximum for the red rays, gives us also the third maximum for the violet rays On the second colour band therefore, the violet of the third band has already encroached, and the red is made less pure As we go further and further from the centre, it is easily seen that the bands of different colours overlap each other more and more The colours we see become, therefore, less and less pure, and very soon overlap sufficiently to form white

An instructive experiment can be made, if a spectroscope is placed in such a position that its slit is parallel to the interference bands, and made to coincide with *e g* the 10th or 20th one of them The appearance in the spectroscope is then represented by equation (3) in which we must now imagine x to be constant and λ to vary

Squaring the expression we obtain for the intensity,

$$4a^2\cos^2\left(\frac{\pi}{\lambda}\frac{xc}{f}\right)$$

The intensity varies periodically, and the spectrum is crossed by bright and dark bands Starting from any one bright or dark band, the next bright or dark band is obtained by increasing $xc/\lambda f$ by unity The bands would therefore be equally spaced in a spectroscope in which the dispersion is proportional to the frequency, *i e* one in which the separation of colours is somewhere intermediate between that of a grating and that of a prism

Writing n for the inverse of λ, so that n is proportional to the frequency, the difference in the values of n_1 and n_2 in two successive bright or dark bands, is given by the relation

$$n_2-n_1=\frac{f}{xc}$$

As an example, we may take the case where the slit of the spectroscope is coincident with the 20th bright band of the red ray whose frequency is defined by N_1 This gives

$$N_1=20\frac{f}{xc},$$

and therefore for the distance between successive bands,

$$n_2-n_1=\tfrac{1}{20}N_1$$

The frequencies of the violet defined by N_2 would be such that

$$N_1 = \tfrac{2}{3} N_2$$

or

$$N_2 - N_1 = \tfrac{1}{2} N_1 = 10\,(n_2 - n_1)$$

The whole range of the spectrum occupies, therefore, ten times the distance between two successive bands, and if there is a bright band in the red, there is also a bright band in the violet, and eight bright bands between them

36. Difficulty of illustrating simple interference phenomena by experiment The simple mathematical treatment of the interference phenomena which we have so far studied, neglects certain effects which disturb the simplicity of the experimental verification Thus the biprism of Fresnel (Fig 38) shows interference only in the angle POQ, but a wave diverging from A and limited at O so that the extreme geometrical ray is AP is not propagated entirely like a complete spherical wave Certain so-called diffraction effects which will have to be discussed in detail, take place, these alter the distribution of light, especially in the neighbourhood of the extreme rays OQ and OP, and there appears a rhythmic variation in the brightness of the fringes, which sometimes makes their measurement difficult The bands seen in Fresnel's mirrors are subject to the same irregularity, owing to the limitation of the beams by the rays AO and BO. Young's arrangement is free from this particular defect, but suffers from another The slits at P and Q do not radiate light equally in all directions, but the intensity is a maximum in the directions OP and OQ respectively, and there are some directions (Art 53) in which the light is totally absent Hence here also, though from a different cause, the experiments give a rhythmic variation in the intensity of the interference fringes, which affects to some extent the positions of the maxima We are therefore led to look in another direction for experimental methods to show interference in its simplest form

37. Light incident on a plane parallel plate. When light is incident on a plane parallel plate, images of the source are formed by reflexion at the two surfaces, the reflected and transmitted beams will then show interference effects due to the overlapping of the waves coming from these images As there are more than two images, the calculations are a little more complicated, but the result being entirely determined by interference, is also more completely represented by the calculation

Let LM and $L'M'$ (Fig. 39) be the parallel surfaces of a transparent plate, and AB an incident plane wave-front, which gives rise to a reflected wave CD and a refracted wave RS This refracted wave will be reflected internally so as to be parallel to $R'S'$ and

however many internal reflexions take place all wave-fronts inside the plate are equally inclined to the surfaces and must be either parallel to RS or to $R'S'$ Similarly all waves which pass out of the surface LM must be parallel to CD and all those passing out of the surface $L'M'$ must be parallel to $A'B'$ which is parallel to AB In calculating the intensity of the waves formed by the combined wave-fronts of the reflected and transmitted light, we shall neglect any absorption of light in the glass plate, so that the sum of the intensities of the reflected and transmitted beams must equal the intensity of the incident beam

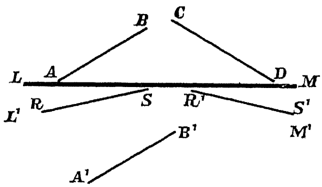

Fig 39

The first step in the calculation must consist in obtaining the differences in phase of the different coincident wave-fronts Making use of the fact that we may calculate optical distances between two wave-fronts along any ray connecting them, we may take some one wave-front AB in the incident beam (Fig 40) and some one ray BS, which at S is reflected towards P Tracing the ray PS backwards through the plate, we find a ray $B'H$ such that starting from the original wave-front AB, it coincides with SP after one internal reflexion at E and refraction at S The phase at P of the wave to which this ray belongs is determined by its phase at B, and the optical length of $B'HESP$ Similarly we may obtain a number of rays through B', B'', B''', B^{iv}, which will coincide along SP having been reflected two, three and four times, at the lower surface of the plate

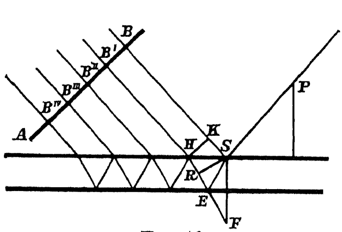

Fig 40

The difference in optical length between the two first rays is the same as the difference in optical length between $HE+ES$ and KS, K being the foot of the perpendicular from H on BS If RS is drawn at right angles to HE, the optical length of KS is the same as that of HR This follows from the fact that HK and RS are parallel respectively to the incident and reflected wave-fronts The difference in optical length is now that due to the path $RE+ES$, or drawing the normal to the plate through S and producing HE to F, its point of intersection with the normal, the difference in optical length is μRF where μ is the refractive index of the plate Noting that SF is twice the thickness of the plate (e), and that the angle at F is the angle between the refracted ray HF and the normal to the plate, for

which we may write γ, it is finally found that the difference in optical length is $2\mu e \cos\gamma$ To obtain the difference in phase at P, we must, however, take account of the fact that the reflexions may be accompanied by change of phase, and we have already shown (Art 25) that according to the principle of reversibility, there must be a change of two right angles at either internal or external reflexion

A difference of phase of two right angles is equivalent to the addition to the optical length of a quantity equal to half the wave-length λ measured in vacuo The difference in phase is therefore finally

$$2\mu e \cos\gamma + \frac{\lambda}{2}$$

If we only considered the superposition of the wave which is reflected externally, and the one reflected once internally, we should find that the intensity of the reflected wave would be zero whenever

$$2\mu e \cos\gamma + \frac{\lambda}{2} = \frac{m\lambda}{2},$$

m being an odd number, or by transposing, when

$$2\mu e \cos\gamma = n\lambda,$$

n being any integer It will be noticed that the difference in path becomes less, as γ, and therefore also the inclination of the incident beam, increases

Before discussing the bearing of this equation, we extend the investigation so as to include multiple reflexions.

We take the vibration at S, due to the incident light, to be represented by $\cos \omega t$, for which, according to Art. 8, we write $e^{\iota\omega t}$, rejecting at the end of the investigation the imaginary part The vibration at S in the reflected wave may then be written $re^{\iota\omega t}$, where r is real, if there is no change of phase An incident wave of unit amplitude would then be reflected as a wave of amplitude r.

We may similarly apply coefficients t to the waves which are transmitted from the outside to the inside, t' for waves transmitted from inside to outside, and s for waves reflected internally A change of phase would be indicated by the coefficients ceasing to be real

Taking account of the fact that each of the rays in Fig (40) has passed through a distance which is longer than the preceding one by the same quantity, of which we have already found the optical equivalent to be $2\mu e \cos\gamma$, the corresponding difference in phase is

$$\delta - 4\pi\mu e \cos\gamma/\lambda$$

Hence we may write for the vibration at S of that ray which has been once reflected internally $stt'e^{\iota(\omega t-\delta)}$, and for the ray reflected internally

three times, the expression becomes $s^3tt'e^{i(\omega t-2\delta)}$, so that the factor $e^{i\omega t}$ in the complete effect at S becomes

$$r + stt'\left(e^{-i\delta} + s^2e^{-2i\delta} + s^4e^{-3i\delta} + \quad\right)$$

The terms of the geometric series in brackets converge towards zer and may be added up We thus find for the amplitude

$$r + stt'\frac{e^{-i\delta}}{1-s^2e^{-i\delta}}$$

Dealing similarly with the transmitted waves, the successive vibratio at E, on emergence, are found to be $tt'e^{i(\omega t-\epsilon)}$, $s^2tt'e^{i(\omega t-\epsilon-\delta)}$ etc if is the difference in phase between S and E, so that the factor of $e^{i(\omega t-}$ in the resultant vibration becomes

$$tt'\left(1 + s^2e^{-i\delta} + s^4e^{-2i\delta} + \quad\right),$$

or

$$tt'\frac{1}{1-s^2e^{-i\delta}}$$

Experiments show that we are justified in assuming that reflexi and refraction at the surface of transparent bodies involve no chan of phase, except, in certain cases, a change of π which can be repr sented by a reversal of sign of the amplitude We may then app the relations found in Art 25 for this case, *i e.*

$$r + s = 0,$$
$$tt' + r^2 = 1$$

The expressions for the reflected and transmitted beams th become for the reflected wave

$$r\frac{1-e^{-i\delta}}{1-r^2e^{-i\delta}}e^{i\omega t}$$

and for the transmitted wave,

$$\frac{1-r^2}{1-r^2e^{-i\delta}}e^{i\left(\omega t-\frac{\delta}{2}\right)}$$

The intensities I_r and I_t for the reflected and transmitted wav are obtained from the results of Art 8 after a simple transformation

$$I_r = \frac{4r^2\sin^2\frac{\delta}{2}}{1+r^4-2r^2\cos\delta} = \frac{4r^2\sin^2\frac{\delta}{2}}{(1-r^2)^2-4r^2\sin^2\frac{\delta}{2}} \qquad (4),$$

$$I_t = \frac{(1-r^2)^2}{1+r^4-2r^2\cos\delta} \qquad (5),$$

from which it follows, as expected, that

$$I_r + I_t = 1$$

The reflected and transmitted beams are therefore always compl mentary In the reflected beams, the intensity is zero whenever

$$\sin\frac{\delta}{2} = 0,$$

i.e when $$2\mu e\cos\gamma = m\lambda \qquad (6)$$

This is identical with the condition for extinction when only two rays were considered, and these were taken to be of equal intensity

The maximum intensity takes place, as is easily seen from (4), when $r \sin \frac{\delta}{2} \Big/ (1 - r^2)$ is a maximum

Both r and δ depend on γ, and a calculation of the condition of maximum intensity in terms of γ would be complicated When the plates are thick, r varies more slowly than δ and we may in that case take as condition for the maximum $\sin \frac{\delta}{2} = 1$, so that $2\mu e \cos \gamma = \frac{m\lambda}{2}$, m now being an odd number We proceed to discuss the special cases to which the above investigation may be applied

38. Colours of thin films. Although there is no difference in principle between the interference effects observed with thin films or with thick plates, the method of observation most favourable in one case is not suitable in the other, and it is therefore convenient to treat each separately. We consider first the case of thin films

The difference in optical length of successive rays ($2\mu e \cos \gamma$) is proportional to the thickness e, and changes much less rapidly with the inclination of the incident beam when the films are thin than when e is large This is especially the case when the incidence is nearly normal so that $\cos \gamma$ is nearly one If in Fig 40 the eye is placed at P, each ray such as SP, forming part of a group coming from a plane wave-front AB, will have the amplitude determined by (4) If the eye is focussed for infinity, we may imagine the whole complex of rays as shown in the figure to be shifted a little to the right or left, and the pupil will collect together a number of these rays parallel to SP, for each of which the amplitude is the same Hence, when the light is homogeneous, the eye will be affected by a reflected beam having the calculated intensity This intensity varies with the inclination of the incident beam, and if the eye looks in different directions, it perceives periodic variations in intensity. Alternate bright and dark bands appear which form part of circular rings, having as centre, N, the foot of the perpendicular drawn from the eye at P, to the surface of the plate

If the source of light is extended, as will generally be the case, a number of mutually inclined wave-fronts fall on the film, and from each of these, those rays may be selected which unite at S An eye focussed on S will combine all these, but if the film is sufficiently thin, the difference in path is approximately the same for the different sets of wave-fronts, and the observed effect is not materially altered.

A peculiarity of thin films lies in the colour effects which are

perceived when the incident light is white The colours of soap bubbles, of thin films of glass, or of a thin layer of oily matter on a sheet of water, are due to this interference of light, reflected from the two surfaces of the film According to the equation (6) any particular wave-length is absent in the reflected beam if

$$2\mu e \cos\gamma = m\lambda.$$

If we imagine the eye to look in one particular direction, a number of different vibrations may be blocked out by interference according to the different integer values of m. When the light is examined by a spectroscope, dark bands will therefore be seen to traverse the spectrum

In order to calculate the number of bands appearing in any portion of the spectrum, we may with sufficient accuracy for our purpose neglect the dispersion in the film, and take μ to be a constant If a wave-length Λ_1 is blocked out, the condition

$$2\mu e \cos\gamma = m\Lambda_1$$

must be satisfied and the Nth band after this must similarly satisfy

$$2\mu e \cos\gamma = (m+N)\,\Lambda_2$$

Eliminating m we obtain

$$N = \left(\frac{1}{\Lambda_2} - \frac{1}{\Lambda_1}\right) 2\mu e \cos\gamma$$

The number of minima included in the interval between a wave length slightly smaller than Λ_2 and one slightly larger than Λ_1 i $N+1$ where an average value of μ may be substituted on the right hand side

If we take for the limits of the visible range the wave-length $6\ 5 \times 10^{-5}$ and 4×10^{-5}, the number of bands included in the rang is found to be nearly

$$2\mu e \cos\gamma \times 10^4$$

For a refractive index equal to that of water, and $\gamma = 45°$, th number becomes $19000\,e$ We can only expect to see colour effec when there are comparatively few wave-lengths blocked out, a otherwise each colour would be weakened in nearly the same pr portion, and the total effect on the eye would again be white not more than three bands are to appear in the visible portion the spectrum, so that $19000\,e$ is not more than 3, e must not l more than $1{\cdot}6 \times 10^{-5}$ This is about double the wave-length at th limits of the red end of the spectrum For bright colours to appe on thin films, the thickness must therefore be of the order of magnitu of a wave-length of light

39. Fringes observed with thick plates. When the plat are thick, nearly homogeneous light must be used Taking *I*

(Fig 40) to be one of the incident rays, it is necessary for complete interference that all the rays passing through B, B', B'', etc should be brought into coincidence, or at any rate, all those which have sufficient amplitude to contribute appreciably to the effect When the plate is thick, these rays may be too wide apart to enter the pupil together, and in that case, observations must be made through a telescope If, on the other hand, SP is considered to be the reflected ray, it follows that the incident wave must have a sufficient width Hence it is necessary for satisfactory observations to have either a wide incident beam, or to collect a sufficient number of rays in the reflected beam

The complete wave-front AB gives rise to a number of rays parallel to SP, which have the same phase along a plane drawn through P and normal to SP The eye should therefore be adjusted for infinity, or a telescope used Fringes produced by the reflexion of light at the two surfaces of a plane parallel plate of thickness considerably greater than the wave-length of light, were first observed by Haidinger, but we owe their investigation more particularly to Mascart* and Lummer† If a plate of glass AB (Fig 41), a few millimetres thick, and carefully worked so as to have its faces plane and parallel, be illuminated by a sodium flame at S, and the rays are reflected from a transparent plate CD, so as to strike the plate AB in a nearly normal direction, the waves partly entering AB and partly reflected at the front surface, will cause overlapping wave-fronts to leave the plate After traversing CD, these rays may be made to enter an eye adjusted for infinity Rings are then observed having for centre the point which is the foot of the perpendicular drawn from the nodal point of the eye to the plate If it is not desired to observe the complete rings, the plate CD may be dispensed with, and the flame placed near the eye If the light reflected from the plate is then directly observed, it is found to be traversed by curved fringes For greater brightness, we may use Lummer's arrangement, in which CD is replaced by a concave perforated mirror such as that used by oculists The rings are observed through the aperture at the centre of the mirror

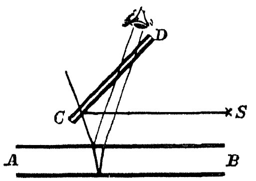

Fig 41.

The condition for extinction is as before

$$2\mu e \cos \gamma = m\lambda,$$

where m is some integer and γ the angle of refraction in the plate

* *Ann Chim Phys* XXIII p 16

† *Wied. Ann* XXIII p. 49.

If θ be the angle of incidence, and h the distance of the eye from the plate, the radius of the ring is

$$h \sin \theta = \mu h \sin \gamma$$

This leads to the following construction (Fig 42) On the radius OA of a circle take a point K_1 such that $\frac{OK_1}{OA} = \frac{m_0 \lambda}{2\mu e}$, where m_0 is the highest integer which gives to the fraction on the right-hand side a value less than one. From K_1 mark off towards O, equidistant points, K_2, K_n, the equal distance being $\lambda OA/2\mu e$ From K_1, K_2, etc draw perpendiculars to OA meeting the circle in S_1, S_2 Then on a scale in which OA represents μh, the diameters of successive dark rings from the centre outwards are given by $K_1 S_1$, $K_2 S_2$, etc

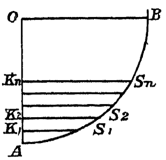

Fig 42

The number of rings is finite, and equal to the highest integer number which is less than OA/K_1K_2 The centre of the system of rings may be bright or dark according to the thickness of the plate and the refrangibility of the light used

Fig 2 Plate I is a photograph of Haidinger's fringes obtained with a glass plate 3 6 mm thick The source of light was a small mercury lamp, the ultra-violet rays being absorbed by a solution of quinine. The actinic light was in consequence very homogeneous, being almost exclusively due to a violet mercury line

A modification of Haidinger's fringes in which multiple reflexions are utilized, has been made use of by Messrs Fabry and Perot in the construction of an apparatus which is equivalent to a spectroscope of very high resolving powers (*Comptes Rendus* CXXVI p 34, 1898)

40. Michelson's combination of mirrors A very powerful optical combination for obtaining interference fringes was devised by Prof Albert Michelson* In principle, it is identical with the system which has just been discussed in the previous article

Let s, Fig 43, be a source of light sending out waves towards a plate of glass a inclined at an angle of 45° to the wave-front The mirror is lightly silvered at the front, so that the incident light is divided into two approximately equal portions, one being transmitted towards a mirror c and the other reflected towards another mirror b At both these points, reflexion takes place which sends the rays back towards a Here once more there is partial reflexion and transmission, and two sets of

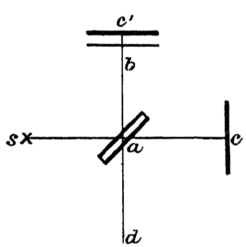

Fig 43

* Described in a joint paper by A A Michelson and Ed W Morley, *Phil Mag* Vol XXIV. p 449 (1887)

wave-fronts will proceed from a towards d, one having passed over the course *sabad*, and the other over the course *sacad*. Neglecting the thickness of the plate a, we may replace the optical length of the path of the ray which passes from a to c and back, by the optical length of the image of that ray from a to c' and back. The interference effects are identical therefore with those obtained when reflexion takes place at two surfaces b and c' of a plane parallel plate.

The additional trouble of adjusting the mirror at c, so that its image is parallel to the mirror b, is counterbalanced by the greater command we have over the different parts of the apparatus. There are also two simplifications in the theory of the observed interference fringes. The passage of the rays taking place in air, the difference in optical length equal to $2 \times bc'$ is the same for all wave-lengths (neglecting the dispersion in air) and there are no multiple reflexions to be considered.

In practice, the thickness of the plate a has to be compensated. This is done by interposing an equal plate f (Fig. 44) into the path of the rays between a and b. The inspection of the figure shows that as one set of waves has been reflected externally at a, while the other has been reflected internally, the retardation in phase affects the results exactly as in the previous cases. We may therefore write down the condition which determines the diameter of a dark ring, by taking $\mu = 1$ in the previous formula, and replacing the angle of refraction γ by the angle of incidence θ on the mirror. Hence

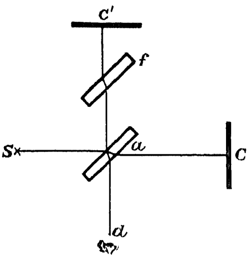

Fig. 44

$$2e \cos \theta = m\lambda$$

is the condition of darkness, m being an integer.

Normal incidence is accompanied by the greatest difference in path, and therefore leads to the highest value of m, which must be that integer which gives to θ its smallest possible value. Calling this m_0 the successive dark rings are obtained by substituting $m_0 - 1$, $m_0 - 2$, etc. for m in the above formula.

As the diameter of the rings is proportional to $\sin \theta$, and θ is so small that powers higher than θ^2 may be neglected, we may write, to this approximation

$$\sin^2 \theta = 2(1 - \cos \theta) = 2\left(1 - \frac{m\lambda}{2e}\right)$$

$$= 2\left(1 - \frac{m_0\lambda}{2e}\right) + \frac{s\lambda}{e},$$

where $m = m_0 - s$. The first term is constant. By giving to s the successive values 1, 2, 3, etc. we obtain the values of $\sin^2 \theta$ for

successive minima If the centre itself is a point of minimum light $\frac{m_0\lambda}{2e}=1$ and the radii of successive dark rings are as the square roots of successive integer numbers, or what comes to the same thing as the square roots of successive *even* integers If the centre is a point of maximum light, we find similarly that successive dark rings are as the square roots of successive *odd* integers

41 Newton's rings. The system of rings observed near the point of contact of a lens, is one of the oldest of known interference phenomena Its elementary theory is simple, though its complete investigation is troublesome, and does not possess sufficient interest to warrant the labour spent on it The colours observed in Newton's rings are the colours of thin films, the film being the layer of air included between the lens and the plate on which the lens is placed. The characteristic distinction between Newton's rings and the phenomena we have already discussed, is that the film has now a variable thickness

The simplest case of a film of variable thickness would be presented by a transparent wedge (Fig 45)

A plane wave AB falling on the wedge, we may select one ray $B'S$ reflected at S towards P, and another $B''H$, such that after refraction at H and reflexion at E, it meets the upper surface at the same point S Owing to the inclination of the two surfaces, the refracted ray SP' is not now coincident with SP, though the inclination is small, if the angle of the wedge is small The difference in optical length between the two rays is $2\mu e\cos\gamma$, or taking account of the change of phase at reflexion, $2\mu e\cos\gamma+\frac{\lambda}{2}$ In this expression,

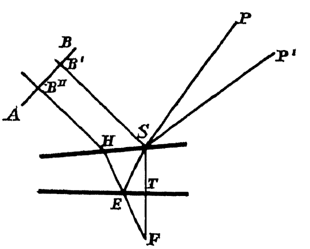

Fig 45

e denotes the length of the perpendicular from S to the lower surface of the plate (which may be taken to be the thickness of the plate at S) and γ is the angle of incidence on the *lower* surface The inspection of the figure explains how the expression is derived Neglecting all rays which have suffered more than one internal reflexion, an eye placed so as to receive both rays SP and SP' and focussed on S, will observe a maximum or minimum of light, according as $2\mu e\cos\gamma$ is an odd or even multiple of half the wave-length If the source of light be extended, waves coming from different directions must be considered Each of these waves supplies two interfering rays at S, and the difference in path depends to some extent on the inclination Hence the eye focussed at S combines on the retina a number of rays which are not under

identical conditions, and the interference will not be so simple or so complete as when the plate has equal thickness throughout The disturbing effect will be small when the plate is thin, and may be neglected for the first few rings in Newton's experiment when the thickness does not exceed more than a few wave-lengths For thicker plates, observations may be improved by reducing the aperture of the pupil by interposing a slit so as to narrow the pencil of light which can enter the eye If the eye is focussed for a point different from S, interference is still observed, though with a slightly changed difference in path This is shown by imagining the ray $B''H$ to be shifted either to the right or to the left If it is shifted to the right, SP' moves to the right, and its intersection with SP moves away from P, so that the eye has now to focus for a more distant point The inclination of SP' remains the same, but the length of the path inside the plate is longer or shorter according as BH'' moves to the right or left It follows from what has been said that we may apply the equations of plane parallel plates to films of varying thickness, so long as their thickness is small The interference is made more complete by restricting the source so that it only subtends a small angle at the film If the incidence is nearly normal, a slight variation in the direction of the incident beam has very little effect on the difference in optical length of the two interfering waves, which also prevents confusion of the interference effects Thus while in the case of the plane parallel plates previously considered, the colours are due to the varying inclination of the incident beam, the thickness of the plate being everywhere the same, we now confine ourselves to the same direction of incidence, and obtain the colours as a consequence of the changing thickness Fig 46 illustrates how Newton's rings may be observed under nearly normal incidence The lens LL' being placed on the plate AB, an inclined transparent plate MN serves to reflect light coming from the source at S, while the eye observes the light reflected from the film included between the lens and the plate, and transmitted through MN To calculate the diameter of the rings, it is only necessary to obtain a relation between the thickness at any point e, and the corresponding distance ρ from the point of contact, if R is the radius of curvature of the lower surface LOL' of the lens, the geometry of the circle gives

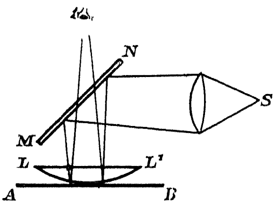

Fig 46

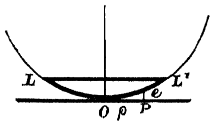

Fig 47

$$\rho^2 = e\,(2R - e)$$

or so long as e is small compared with R, so that its square may be neglected,

$$\rho^2 = 2eR$$

The difference in optical length of the two interfering rays is $2e + \frac{\lambda}{2}$, if the observation is conducted so that the medium included between the lens and plate is air, so that $\mu = 1$ The diameters of the rings of maximum illumination are obtained by making $2e$ an odd multiple of half a wave-length, so that they are proportional to the square roots of successive *odd* numbers, while the dark rings will have diameters proportional to the square roots of successive *even* numbers The centre of the ring system is dark, though black only when the lens and plate are of the same material, in which case the whole light is transmitted if there is optical contact at O

The minimum of light at the centre of the system of rings appears as a consequence of the retardation $\lambda/2$ at internal or external reflexion, there being in consequence total destruction with no difference of path If the upper and lower surfaces are made of different material, and the film has a refractive index intermediate between the two, the centre of the ring system on the other hand is bright, as the half wave retardation now disappears Thomas Young showed this by introducing oil of sassafras between a lens of crown glass and a plate of flint glass

In the transmitted system of Newton's rings, the colours are less brilliant Their position is easily deduced from the fact that the effect at every point must be complementary

Plate II Figs 3 and 4 represent photographs of Newton's rings. The same mercury lamp was used in both cases as the source of light, the rays producing the photographic effect being principally derived from one violet and one ultra-violet radiation In Fig 3, the ultra-violet radiation has been blocked out by an absorbing screen and hence the appearance is that due to practically homogeneous light In Fig 4 we may observe the effect of the overlapping of two systems of rings which alternately strengthen and neutralize each other Where the dark and bright rings of the two systems coincide the rings are clearly defined, where the bright ring of one overlaps the dark ring of the other the rings are very indistinct Similar effects may be observed with sodium light owing to the difference in the wave-lengths of the two components of the sodium doublet, but the two wave-lengths being more nearly equal the intervals between the regions of maximum definition are much greater

42. Brewster's bands When light passes through a plate of glass, a small change in optical length may be made by slightly inclining the plate. If it is desired to observe interference effects due

PLATE I

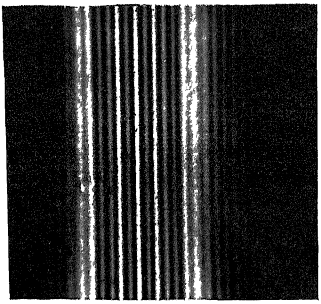

FIG. 1

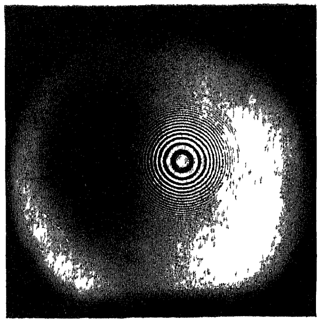

FIG. 2

FIG. 3

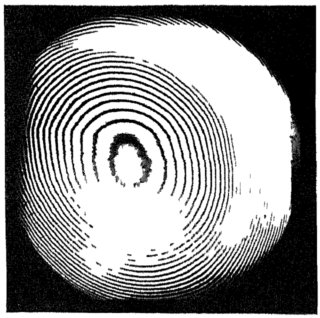

FIG. 4

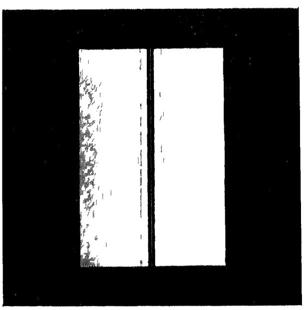

FIG. 5

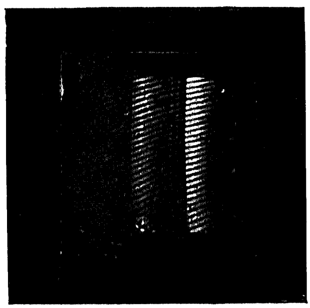

FIG. 6

[To face page 74

to this alteration, it is necessary to interpose an exactly equal plate into another portion of the same beam, so that in the first instance there may be equality of path, which is then slightly disturbed by the inclination of one of the plates This leads to the following arrangement due to Brewster A plate of glass, which should be as nearly as possible plane parallel, is cut in half so as to obtain two plates of equal thickness, one is slightly inclined to the other and light passed through them The course of the rays which are brought to interference is shown in Fig 48 If the plates were parallel, the optical lengths would be equal A slight inclination of one causes a relative change of phase in the overlapping beams, and when an illuminated surface is viewed through the plates, coloured bands are seen to traverse the field The interference fringes may also be observed in reflected light, and Fig 49 shows how we may obtain a number of different sets of interfering rays according to the number of internal reflexions. In the first system, marked 1 in the figure, two rays are brought to interference, the first having been once reflected internally in the plate A, and the second once in plate B The second system consists of three rays, one of which has been reflected once in each plate, and the two others twice in one plate and not at all in the other The course of the third system is also shown in the figure, and the further ones need hardly be considered, as the intensity of light rapidly diminishes by multiple reflexions To prevent confusion, it is necessary to place a screen at SS' to limit the incident beam If the bands were observed near the plane of the figure they would be seen to be strongly curved, and the field of view would only contain bands formed by rays having large retardations To find the position of the central band which is that in which the relative retardation is zero, we start from the fact that the optical length in each plate depends only on the angle of incidence of the light. The thickness and refractive index of the two plates being the same, the optical length is the same for all rays which in their passage from one plate to another are equally inclined to both plates. These rays all lie in a plane which is parallel to the line of intersection of the plates and equally inclined to both of them The image of that plane in the plate B is the locus of the central band. If we wish to carry on our observations in a horizontal plane, the two

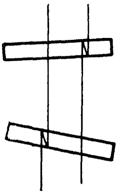

Fig 48

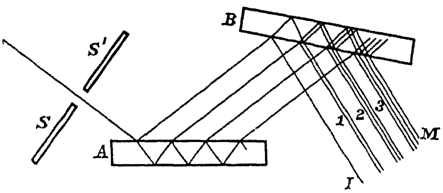

Fig 49

plates should therefore be equally inclined to the horizon through a small angle but in opposite directions, their intersections being horizontal. If the second plate is inclined downwards the central fringe will lie a little below the horizontal plane The fringes near the central plane are approximately horizontal, and if a spectroscope is used to analyse the reflected light, the slit should be horizontal. Useful hints for mounting the apparatus in this and other experiments on interference are given in Quincke's writings*

43. Stationary vibrations. When two waves of the same amplitude and period proceed in opposite directions, we may represent the displacement by

$$a \cos (\omega t - nx) + a \cos (\omega t + nx) = 2a \cos nx \cos \omega t.$$

The right-hand side of the equation shews that the phase is now constant everywhere, but the amplitude depends on x and is zero whenever x is an odd multiple of a quarter of a wave-length. The amplitude has a maximum value in the intermediate places at which x is a multiple of half a wave-length The combined disturbance of two waves crossing each other in this way is called a stationary vibration An alteration in the phase of one of the waves shifts the positions of the maxima and minima, but does not alter their distance. Altering *e g* the phase of the wave proceeding in the negative direction, by two right angles, we should get

$$a \cos (\omega t - nx) - a \cos (\omega t + nx) = 2a \sin \omega t \sin nx$$

These stationary waves are easily illustrated in the case of sound waves Experimental investigation in the case of light involves great difficulties, which were, however, successfully overcome by O Wiener†. His experiments required the preparation of a photographic film having a thickness considerably smaller than a wave-length of light. They were made of collodion, sensitized by chloride of silver, and had a thickness from between one-twentieth to one-fortieth of a wave-length. Light was incident normally on the silver coating of a plate of glass, and the sensitized layer of collodion was placed against this plate, in such a fashion that when the bands of thin films appeared in reflected light, they were several millimetres apart This means that the film formed a very small angle with the silver surface, it being a necessary condition for the success of the experiment, that the two surfaces should not be quite, though very nearly, parallel

When nearly homogeneous light is made to fall normally on this arrangement, the stationary vibrations give rise to disturbances which are of different amplitude at different distances from the reflecting

* *Pogg Ann* Vol 132, p 29 (1867)

† *Wied. Ann.* XL p. 203 (1890)

surface. As the film is slightly inclined the photograph when developed is crossed by alternately bright and dark bands, the dark bands being due to the deposit of silver at the places where the amplitude was near its maximum Plate II Fig 5 is a reproduction of one of Wiener's photographs The light incident on the film was decomposed in this case into its homogeneous constituents by a prism, and the result gives a picture of the separate effects of the different wave-lengths The vertical bands represent the spectrum of the electric arc used The carbon bands will be noticed, and the *H* and *K* calcium lines show faintly, but each of these lines and bands is seen to be crossed obliquely by a series of bands which are due to the variations in amplitude of the stationary vibration The inclination of the two systems of bands to each other is due to the inclination of the refracting edge of the prism decomposing the light to the edge of the wedge formed by the photographic film and the reflecting plate

Drude and Nernst having succeeded in obtaining sufficiently thin fluorescent films, observed the stationary vibrations by their fluorescent effect

Lippmann's Colour Photography is based on the formation of thin layers of reduced silver deposited within a photographic film, the layers being half a wave-length apart They are formed by the stationary vibration of waves of light reflected from a surface of mercury over which the sensitive film has been extended When viewed in reflected light, the colours of thin plates are seen, and that colour shows a maximum of intensity which has a wave-length equal to twice the distance between the layers We therefore see in the reflected light chiefly the colour belonging to the wave which originally had formed the stationary vibration The possibility of reproducing natural colours by photography in this fashion, had already occurred to Wm Zenker*, and to Lord Rayleigh† The experimental realization due to Lippmann is, however, a very considerable experimental achievement

44. Applications. We may divide the principal applications of the interference phenomena which have been described, into two classes In the first, the difference in phase of two portions of the same wave-front which have traversed two media, is used to measure the difference in optical length of the path, and hence the difference in refractive index Instruments used for this purpose have been called interference refractometers Fresnel, already, in conjunction with Arago, used the principle of interference to measure the difference in the refractive index between dry and moist air Two parallel tubes, filled with the gases to be examined, were placed in the path of a plane wave-front

* *Lehrbuch der Photochromie*

† *Collected Works*, Vol III p 13

which traversed the tubes longitudinally. the displacement in the bands observed when dry air was replaced by moist air served as a measure of the difference in refractive index. Jamin carried out important measurements in an apparatus in which use was made of Brewster's interference bands. Fig. 50 represents Jamin's apparatus.

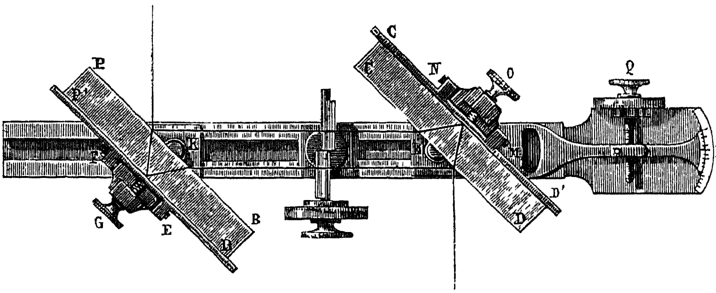

Fig. 50.

One of the plates of glass PB is fixed, while the other CD is movable round a vertical axis by means of the screw Q and a horizontal axis by means of the screw O. If the plates are exactly parallel, the illumination over the whole field is uniform. If now the second plate be slightly tilted round a horizontal axis, horizontal bands appear which come closer and closer together as the inclination increases. Rotation round the vertical axis does not change the distance between the bands, but shifts the whole system of bands up and down. If two tubes, which may be exhausted or filled with different gases, be introduced into the path of the interfering rays, the position of the bands depends on the relative retardation of the two rays. Starting *e.g.* with air at atmospheric pressure, and the central band being adjusted so as to pass through the centre of the field, a partial exhaustion of one of the tubes displaces the bands, the displacement being measured by means of a "Compensator." This compensator, in Jamin's apparatus, consisted of two plates of glass (Fig. 51) capable of being rotated round a horizontal axis AB, and placed at such a distance from each other that each plate receives the light which has passed through one of the tubes. Rotation round the horizontal axis alters the thickness of glass traversed. The alteration being different for the two plates a measurable retardation of one set of rays, as compared with the other, is produced. If the central band, having been displaced by the change of pressure in the tube, is brought back by the compensator to its original position, the difference in refractive index between the air under partial exhaustion and the air at atmo-

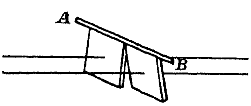

Fig. 51.

spheric pressure, can be measured Different gases may be compared in a similar manner.

Lord Rayleigh's form of Refractometer more nearly approaches the original instrument of Fresnel and Arago

Light coming from a fine slit L and rendered parallel by a collimator lens C of 3 cms aperture passes through two brass tubes side by side, and soldered together

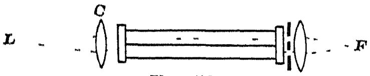

Fig 52

These tubes, 20 cms long and 6 mms in bore, are closed at the ends by plates of worked glass, so connected as to obstruct as little as possible the passage of light immediately over the tubes The light having passed through the tubes enters two slits and is brought to a focus F by means of a lens The optical arrangement is practically identical with that which gives Young's fringes (Art 32) The fringes are observed by means of an eyepiece. To secure better illumination and sufficient magnifying power, the eyepiece is cylindrical, so as only to magnify in a horizontal direction It is made of a short length of glass rod, 4 mm in diameter. There are two systems of bands, one formed by light which has traversed the gases within the tubes, the other by light which passes above them If different gases are to be compared with each other, as regards their refracting power, their pressure is adjusted until the system of bands formed by light which has passed through the tubes is coincident with the system formed by the light which has passed above the tubes, the retardation in the two tubes is then the same If the experiment be repeated at a different pressure, then the ratio of the changes of pressure for each gas is the inverse ratio of the refractivities ($\mu - 1$) of the gases

Other refractometers have been constructed, chiefly with a view to separating the path of the interfering rays laterally as much as possible, so as to leave more room for the tubes or other apparatus to be introduced into the path of the rays It is sufficient to refer to the apparatus of Zehnder* It should be noticed, however, that the lateral separation of the rays is by no means always an advantage One of the experimental difficulties in delicate optical measurements consists in keeping the temperature sufficiently constant, or at any rate, not to introduce a *difference* in temperature into the two optical paths The nearer these are together, the easier will equality of temperature be secured Where a separation of rays is necessary or advisable for other reasons, Michelson's arrangement, which has already been described, will probably be found to be the most advantageous. The applications which Michelson has made with this

* *Ztsch f. Instrumentenkunde*, 1891, p 275

instrument to the investigation of the constitution of nearly homogeneous radiation will be referred to in Chapter XIV

An appliance, useful in many optical measurements, is the "bi-plate" which serves either to separate or to bring together two parallel beams of light It consists of two plane parallel plates of glass cemented together at an angle Their action is sufficiently illustrated by of Fig 53

In the applications of the phenomena of interference which have been dealt with so far, the problems are of a purely optical nature We turn now to the second class of applications in which optical methods are used for linear measurement

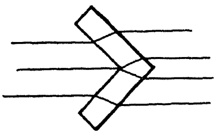

Fig 53

Fizeau has used Newton's rings to examine the coefficients of expansion of certain substances The body to be examined, cut *e g* into the form of a cube, is placed on a plate which, by means of screws passing through it, supports a lens The upper surface of the cube is polished If the lens be adjusted so as to leave a small air space between it and the cube, Newton's rings may be observed If the whole arrangement is raised in temperature, a change takes place in the rings which depends on the altered distance, between the upper surface of the cube and the lens Knowing the effect of temperature on the refractive index of air and the coefficient of dilatation of the other part of the apparatus, that of the cube may be deduced Fizeau has measured in this manner, the expansion of crystals in different directions For a more detailed account of the apparatus and method of obtaining the result from the observed displacements of Newton's rings, Mascart's *Optics*, vol 1, p 503, may be consulted

Perfectly flat surfaces are sometimes required in optical investigations, and it is a matter of great difficulty to work them so as to satisfy optical tests Not the least of the difficulties consists in testing the surface when it is nearly flat, so as to discover where its faults are and how they may be corrected Lord Rayleigh* uses for this purpose the interference bands seen between a horizontal surface of water and the carefully levelled surface which is to be examined The latter surface is supported horizontally at a distance of about one or two millimetres below that of the water. By the aid of screws the glass surface is brought into approximate parallelism with the water When the surface is perfectly flat, the interference bands are straight, while a curvature of the bands always implies a curvature of the surface In the paper referred to it is shown

* *Collected Works*, vol IV p 202

how to interpret the curvature of the surface by means of that of the bands. The chief difficulty in applying the method consists in securing perfect steadiness, so as to avoid the effects of the tremor of the water surface.

An important application of interference bands has been made by Michelson*, who was able to obtain a direct comparison between the standard of length, and the wave-length.

45. Historical. Christian Huygens, (born April 14, 1629, at Haag in Holland, died June 8, 1695,) is the founder of the undulatory theory of light. His treatise on light appeared in 1690, and contains the explanation of the reflexion and refraction of light by means of the principle which now bears his name. He also demonstrated how double refraction could be explained by means of wave-surfaces having two sheets, and in particular showed how, in Iceland Spar, a wave-surface consisting of a sphere and spheroid accounted for the laws of refraction of both rays.

Sir Isaac Newton (born Jan. 5, 1643, in Lincolnshire, died March 21, 1727) did not look with favour on the undulatory theory of light. He was misled by the apparent difference in the behaviour of waves of sound which, after passing through an opening, spread out in all directions, and the rays of light which pass in nearly straight lines. This seemed a formidable difficulty, and Huygens' attempts at explaining the apparently rectilinear propagation of light were not clear or convincing. While there is no doubt that Newton's great authority kept back the progress of the undulatory theory for more than a century, this is more than compensated by the fact that the science of Optics owes the scientific foundation of its experimental investigation in great part to him. His experiments on the prismatic decomposition of white light do not fall within the range of this volume, but the phenomena of Newton's rings have been referred to. Newton discovered that the radii of bright or dark rings was determined by the thickness of the layer of air interposed, and found the correct law connecting the diameters of successive rings.

Thomas Young, born June 13, 1773, at Milverton (Somerset), studied medicine in London, Edinburgh and Gottingen. He was Professor of Physics at the Royal Institution in London between 1801 and 1804, but gave up his position in order to devote himself to the practice of medicine. He died on May 10, 1829. To Young belongs the merit of having been the first to state clearly the principle of the superposition of waves and to show how interference may be explained by means of it. Owing to the historical importance of this

* *Travaux et Mémoires du Bureau International des Poids et Mesures* XI. (1895).

principle, on which the development of the undulatory theory of light entirely depends, the passage in which Young first introduced it may be quoted It occurs in a paper read before the Royal Society on November 12, 1801, and runs as follows

"PROPOSITION VIII When two Undulations, from different Origins, coincide either perfectly or very nearly in Direction, their joint effect is a Combination of the Motions belonging to each"

"Since every particle of the medium is affected by each undulation, wherever the directions coincide, the undulations can proceed no otherwise, than by uniting their motions, so that the joint motion may be the sum or difference of the separate motions, accordingly as similar or dissimilar parts of the undulations are coincident."

Young's arrangement for observing interference fringes, which has been discussed in Art 32, is thus described in his published lectures (1807)

"In order that the effects of two portions of light may be thus combined, it is necessary that they be derived from the same origin, and that they arrive at the same point by different paths, in directions not much deviating from each other This deviation may be produced in one or both of the portions by diffraction, by reflection, by refraction, or by any of these effects combined, but the simplest case appears to be, when a beam of homogeneous light falls on a screen in which there are two very small holes or slits, which may be considered as centres of divergence, from whence the light is diffracted in every direction. In this case, when the two newly formed beams are received on a surface placed so as to intercept them, their light is divided by dark stripes into portions nearly equal, but becoming wider as the surface is more remote from the apertures, so as to subtend very nearly equal angles from the apertures at all distances, and wider also in the same proportion as the apertures are closer to each other The middle of the two portions is always light, and the bright stripes on each side are at such distances, that the light, coming to them from one of the apertures, must have passed through a longer space than that which comes from the other, by an interval which is equal to the breadth of one, two, three, or more of the supposed undulations, while the intervening dark spaces correspond to a difference of half a supposed undulation, of one and a half, of two and a half, or more"

There is no other reference to these experiments in Young's published paper, so that we do not know the size of the apertures, or their distance apart Young seems to have attached more importance to the cases where the openings are wide and the intervening space narrow, though the theory of these cases is more complicated.

Young was very successful in his explanation of the colour of thin films, especially in the mechanical analogy which he brought to bear on the change of phase which takes place at one of the reflexions The otherwise formidable difficulty of explaining the dark centre of Newton's rings was thus at once satisfactorily overcome

Fresnel's important work belongs more particularly to the next chapter As regards simple interference and its experimental illustration, we owe him the method of inclined mirrors and of the biprism. He also showed how fringes could be observed subjectively through an eyepiece, a method of observation which enabled him to carry out accurate measurements

CHAPTER V.

THE DIFFRACTION OF LIGHT

46. Huygens' principle By means of Huygens' principle, we may obtain the effect of a wave-front WF at a point P (Fig 54), by dividing its surface into a number of elements, and adding up their effects Our problem then consists in finding the law according to which a small portion of the surface may be supposed to act If we consider the element at S to be an independently vibrating source, it is seen that its effect at P can only depend on the length of the vector SP, the angle which that vector forms with the normal to the surface, and the angle between the same vector and the direction of vibration at S If the investigation be limited to homogeneous vibrations, we may obtain in a simple manner an expression for the displacement at P which yields, at any rate, one possible solution of the problem

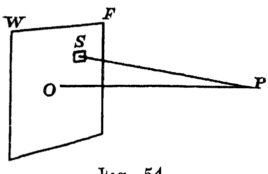

Fig 54

Draw PO, the normal to the wave-front WF, and call O the pole of P Draw two circles with the pole as centre and radii OS and OR The area of the ring included between these circles is

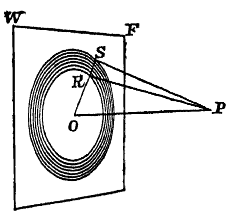

Fig 55

$$\pi (OS^2 - OR^2) = \pi (PS^2 - PR^2)$$
$$= \pi (PS - PR)(PS + PR)$$

If $PS - PR = \delta$, and δ is a small quantity, the square of which can be neglected compared to PO, the expression for the area of the ring becomes

$$2\pi\delta PR$$

If the ring be further subdivided into a very large number of concentric circles having radii OR_1, OR_2, such that

$$PR_1 - PR = PR_2 - PR_1 = PR_3 - PR_2 = \text{etc},$$

the successive rings have equal areas, and their separate effects at P must be equal in magnitude To calculate their joint effect, we must

take account of the difference in phase of the vibrations reaching P from each separate ring If a diagram be drawn in which the effect of each ring is represented by a vector, these vectors will be of equal length and will succeed each other at equal angular intervals

Hence according to Art 5 the resultant amplitude is $a\dfrac{n \sin \alpha}{\alpha}$, where α is half the difference in phase between the first and last vibration The product na represents what would be the amplitude at P, if every point of the ring were at the same distance from that point Writing c for this product, we find that the complete ring causes a vibration at P, having an amplitude $\dfrac{\sin \alpha}{\alpha}c$ Its phase is the arithmetic mean between the phases due to the first and last ring If $PS - PR = \lambda$, $\alpha = \pi$ and the amplitude at P is zero If $PS - PR = \lambda/2$, the amplitude at P is $2c/\pi$

Divide now the whole wave-front into zones, Fig 56, by rings of radii OK_1, OK_2 such that

$$\frac{\lambda}{2} = PK_1 - PO = PK_2 - PK_1 = PK_3 - PK_2 = \text{etc}$$

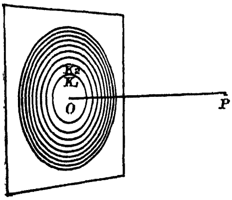

Fig 56

The resultant phase of two successive zones differs according to the above by two right angles, so that to obtain the total effects, we need only add up the effects of successive zones, giving the opposite sign to successive values

Hence $$S = m_1 - m_2 + m_3$$

where $m_1 = 2c_1/\pi$, $m_2 = 2c_2/\pi$ etc The quantities c_1, c_2 etc depend on the distance between each ring and P, and may also depend on the angle KPO or the angle between the direction of vibration and KP. These quantities all alter very little between one ring and the next and we may therefore take the difference between two successive values of m to be very small This being so, a very simple expression for the sum of the series may be obtained

Collecting the terms differently, the series may be written in the form

$$S = \frac{m_1}{2} + \left(\frac{m_1}{2} - m_2 + \frac{m_3}{2}\right) + \left(\frac{m_3}{2} - m_4 + \frac{m_5}{2}\right) + \quad \ldots \quad (1),$$

the last term being $\frac{1}{2}m_n$ or $\frac{1}{2}m_{n-1} - m_n$, according as n is odd or even. Each of the bracketed terms is small if the values of m alter slowly, but we should not be justified for this reason alone in neglecting them, because if their number is large, their sum may be comparable in magnitude to m_1 But assuming that the law of vibration is such that the effect of each zone is smaller than the arithmetic mean of the

effects of the preceding and following zones, all terms in brackets are positive, and therefore

$$S > \frac{m_1}{2} \pm \frac{m_n}{2} \qquad (2),$$

where the plus or minus sign is chosen according as n is odd or even

The series S can also be written in the form

$$S = m_1 - \frac{m_2}{2} - \left(\frac{m_2}{2} - m_3 + \frac{m_4}{2}\right) - \left(\frac{m_4}{2} - m_5 + \frac{m_6}{2}\right) + \qquad (3),$$

and under the same conditions as before, each bracket is positive. If m_1 is sensibly equal to m_2, and m_n sensibly equal to m_{n-1}, it follows that

$$S < \frac{m_1}{2} \pm \frac{m_n}{2}.$$

Comparing this with (2) it is seen that the bracketed terms are negligible, and hence

$$S = \frac{m_1}{2} \pm \frac{m_n}{2}$$

The same conclusion is arrived at by supposing that the brackets in (1) and (3) are all negative If a change in the sign of the brackets occurs in the course of the series, we may divide the series into two parts, and sum each part separately We thus arrive again at the same conclusion that the whole effect is equal to one-half the sum of the effects of the first and last zones, unless the brackets in the expression (2) change so frequently in sign that the outstanding small effect at each reversal sum up to be an appreciable quantity

Excluding such special cases, which need not be considered in any optical application, we may now apply our result to the calculation of the resultant effect of a plane wave-front extended but ultimately limited by a boundary which is not a circle having the pole as centre In Fig 57 the boundary is assumed to be square We may draw all the circles complete until one of them touches the boundary After that point is reached, parts of the zones are blocked out by the opaque screen, and the effect of these outer zones must gradually diminish and ultimately vanish In this case, therefore, the effect of the last zone is zero, and we find that the resultant effect at P is equal in magnitude to half that of the first zone. Writing p for OP, the area of the central disc has been shown to be $\pi p\lambda$, and its effect equal to that of a surface $2p\lambda$ placed at a distance half way between the centre and the edge of the central disc To obtain its effect we must apply the factor $2/\pi$ and thus find that it causes an amplitude at P

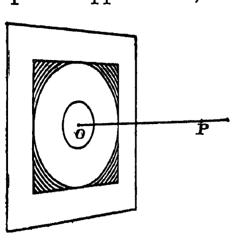

Fig 57

which is equal to a surface of area $2p\lambda$, all points of which are at the same distance from P. If ks is the effect at P of a small surface s placed at O, the effect of the first zone is $2kp\lambda$, and the effect of the whole wave, as has been shown, is equal to that of half the first zone. The wave being plane, the amplitude at P is the same as at O. Calling that amplitude a, $kp\lambda = a$, and hence

$$k = \frac{a}{p\lambda}.$$

We must conclude that if a wave-front is split up into a number of small elements, we arrive at a correct result in the case of an extended plane wave of amplitude a, if we take the effect at a point P of a small surface s as regards amplitude to be $as/p\lambda$. The surface s is here supposed to be so small that the distances of its various points from P do not differ by more than a small fraction of the wave-length. The occurrence of p in the denominator can readily be understood, as the effect of an independent source on a point at a distance may be expected to be such that the intensity varies inversely as the square of the distance. If this be granted, it also follows that λ must occur in the denominator, as the factor of a must be of the dimensions of a number, and of the three quantities s, p, λ involving the unit of length s occurs in the numerator and p in the denominator.

It may now be shown that the value of k just obtained also gives correct results, when the wave-front is spherical. In Fig. 58 let waves diverge from a point Q and let it be required to calculate the effect at P from one of the wave-fronts WF. The only difference there can be between this problem and the previous one lies in the magnitude of the first zone, which must therefore be recalculated.

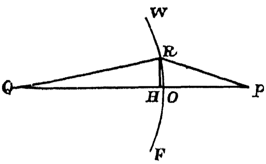

Fig. 58

Let RH be drawn at right angles to PQ and let $QO = q$; $PO = p$, $RH = f$, $HO = t$.

Then neglecting powers of f higher than f^2

$$f^2 = 2qt,$$

also

$$f^2 = PR^2 - PH^2$$
$$= (PR + PH)(PR - PH),$$

and

$$PR = p + \frac{\lambda}{2},$$

$$f^2 = \left(\frac{\lambda}{2} - t\right) 2p.$$

Eliminating t,

$$f^2 = \frac{\lambda pq}{p+q}.$$

The effect of the first zone as regards amplitude is equal to $2ks/\pi$, where s is the surface of the zone. Substituting $s=\pi f^2$ and $k=a/p\lambda$, where a is the amplitude at O, the amplitude at P which is half the effect of the first zone is found to be $aq/(p+q)$ and varies therefore, as it should do, inversely as the distance from Q

Returning to the case of plane waves we obtain another important result by considering the phase of the resultant vibration The phase at P due to the action of any zone has been shown above to be half way between the phases due to portions of the zone which are respectively nearest to and furthest from P Applying this to the central zone, the phase of the resultant vibration at P, if calculated in the usual way, should differ from that at O by

$$2\pi\left(p+\frac{\lambda}{4}\right)\Big/\lambda \text{ or } \frac{2\pi p}{\lambda}+\frac{\pi}{2}$$

But we know that the optical distance from O to P is simply p, and hence the difference in phase is $2\pi p/\lambda$ It follows that if we want to obtain the phase correctly at P by means of Huygens' principle, we must everywhere subtract a quarter of a wave-length from the optical distance, or imagine the wave-front to be shifted forward through that distance

It should be clearly understood what it is that has been proved. An extended wave-front has been divided into zones, and grouped together in such a way that the effect of the whole wave was found equal to that of half the central zone which lies close to the pole O The effect of a small surface s at a distant point P being expressed by ks, the result has shown that the possible variations of k depending on the angle between the normal to the wave-front and the radius vector, do not enter into the question at all We conclude that those portions of the effect which *might* depend on it, are eliminated by interference Similarly the result is independent of any possible effect of the direction of vibration

The division of the wave-front into zones, drawn so that the distance of their successive edges from the point at which the amplitude of light is to be estimated, increases by half a wave-length, has rendered it possible to apply Huygens' principle in a simple and effective way. This mode of treating the propagation of waves being due to Fresnel, the zones should be called "Fresnel Zones"

47 Laminar zones. Instead of dividing the wave-front into circular zones, it is often more convenient to perform the division in a different manner Let P (Fig 59) be the point at which the light is to be estimated and WF the wave-front Divide WF into a number of parallel strips at right angles to a central line HK Let LM be

such a strip, which may again be subdivided into smaller areas, chosen to be of such magnitudes that the resultant phases of two successive elementary areas are in opposite directions If the strip be indefinitely extended in both directions, we may form a series as in the previous article, and find in this way that the total effect must be some definite fraction of that element of LM which is nearest to the central line HK. The whole effect being proportional to the width of the strip t, we may put it equal to kht, where k is the factor previously determined, and h some linear quantity This expression asserts nothing more than that the effect of the strip is equal to that of an area situated in the central line HK, having a width t and a height h The same reasoning may be applied to each of the strips which are parallel to LM, and we finally reduce the effect of the wave-front to that of a horizontal strip of width h This may once more be subdivided As the strip of width t produces an effect at P equal to kht, the effect of a strip of width h must be kh^2 Hence the effect of the complete wave-front is reduced to that of an area h^2 placed at O, O being the pole of P If the amplitude is a, $kh^2 = a$ Hence

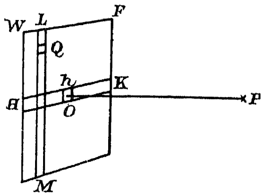

Fig 59

$$h = \sqrt{\frac{a}{k}} = \sqrt{p\lambda},$$

p being the distance OP, the effect as regards amplitude of a strip such as LM of width t is therefore $ta/\sqrt{p\lambda}$

To obtain the resultant phase due to each strip, we make use of the previously established fact that in applying Huygens' principle, we obtain the optical distance by taking away a quarter of a wave-length from the actual distance between the source and the point at which the amplitude is required. We imagine therefore the whole wave-front to be brought nearer through that distance Now the process of attaining the final resultant from the rectangular strips consists of two exactly equal steps, the first in obtaining the intermediate resultant of each vertical strip such as LM, and the second in summing up for the horizontal strip HK which represents that intermediate resultant. If the total effect of the two steps as regards phase, is to bring back the wave-front to its proper position, each step must contribute equally, and therefore the optical distance of each strip is obtained by taking away $\lambda/8$ from the actual distance When the wave-front is divided into strips, it follows therefore that for the calculation of phases, we must imagine each strip to be brought nearer by $\lambda/8$ Or for simplicity of calculation we may say that we may take the optical distance of a strip to be equal to its actual distance, if we correct the final result

by subtracting $\lambda/8$ from the calculated optical distance or 45° from the calculated phase

We may now determine the widths, t, of the strips, so that their resultant effects at some given point are alternately in opposite directions. Let Q be the point and OT_1, T_1T_2, T_2T_3, etc (Fig 60) represent the widths A vertical strip T_nT_{n+1} causes a resultant effect at Q which unless the strip is close to O must be the same as if it were concentrated into a line halfway between T_n and T_{n+1} The total resultant effect of all the vertical strips has been shown to give an optical distance of $p+\frac{\lambda}{8}$ instead of p, and if the resultant phases of successive strips are in opposite directions, the optical distance of the centre of each must be

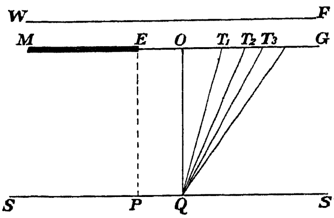

Fig 60

$$p+\frac{4n+1}{8}\lambda$$

This gives for the edges of the strips the equation

$$QT_n=p+\frac{4n-1}{8}\lambda$$

The central strip wants special consideration It would not be correct to say that its resultant phase is the arithmetic mean between that due to the vibrations at O and T_1, its nearest and furthest points, because the distance from Q to the line HK passes through a minimum at O Hence the phase at Q due to any vertical subdivision of the strip, does not alter uniformly with the distance of that subdivision from O It is found, however, that the error introduced by making the supposition for the *second* strip is already very small, and hence the above subdivision will give sufficiently nearly the dividing lines between the zones which yield alternately opposite phases at P, because if the sum of all the strips above the first gives a phase equal to that of the resultant of all the strips, including the first, the phase of the first strip must be opposite to that of the sum of the remaining ones, which is equal to that of the second strip

We take in accordance with this argument

$$QT_1=p+\tfrac{3}{8}\lambda,$$
$$QT_2=p+\tfrac{7}{8}\lambda,$$
$$QT_3=p+\tfrac{11}{8}\lambda \text{ etc.}$$

A more complicated calculation gives as a further approximation

$$QT_1=p+\tfrac{3}{8}\lambda-\ 0046\ \lambda,$$
$$QT_2=p+\tfrac{7}{8}\lambda+\ 0016\ \lambda,$$

showing that for nearly all purposes, the error introduced by the simplification we have made is negligible

The width of successive strips is obtained from

$$OT_n = \sqrt{QT_n^2 - p^2} = \sqrt{\frac{4n-1}{4} p\lambda},$$

where λ^2 is neglected compared to $p\lambda$ Hence for the first strip

$$t_1 = \tfrac{1}{2}\sqrt{p\lambda}\,\sqrt{3},$$

for the second strip $t_2 = \tfrac{1}{2}\sqrt{p\lambda}\,\{\sqrt{7} - \sqrt{3}\}$,

and generally

$$t_n = OT_n - OT_{n-1}$$
$$= \tfrac{1}{2}\sqrt{p\lambda}\,\{\sqrt{4n-1} - \sqrt{4n-5}\}$$

The effect of the nth strip t_n is, as regards amplitude :

$$\frac{2}{\pi}\frac{t_n}{\sqrt{p\lambda}} = \frac{1}{\pi}\{\sqrt{4n-1} - \sqrt{4n-5}\}$$

The numerical values of the effects are given in Table III for $n=2$ to $n=12$ They have been calculated from the above expression, except for the first strip, for which the method fails to give correct results The effect of this strip may be obtained by calculating the numerical value to which the series approaches, leaving out the first strip

The series to be summed up is

$$\frac{1}{\pi}[(\sqrt{7} - \sqrt{3}) - (\sqrt{11} - \sqrt{7}) + (\sqrt{15} - \sqrt{11}) \qquad]$$

Its value is found to be 1725, and the effect of all strips on one side of O being 5, if the amplitude of the incident wave is unity, it follows that the first strip produces an effect equal to 6725, as it is in the opposite direction as regards phase to the resultant effect of the rest of the wave-front

TABLE III

Effects in Amplitude of Fresnel Strips

No of strip	Effect in amplitude	No of strip	Effect in amplitude
1	+ 6725	2	− 2908
3	2135	4	1771
5	1547	6	1391
7	1274	8	1183
9	1109	10	1047
11	0995	12	0949

48. Preliminary discussion of problems in diffraction. When an obstacle is placed in the path of a wave-front and the shadow of the obstacle received on a screen, the boundary of the shadow is not sharp, but the light encroaches to some extent on the dark portions, while there are bright and dark fringes on the side towards the light. If we draw straight lines which proceeding from the source of light touch the shadow-throwing body, the intersections of these lines with the screen enclose what may be called the geometrical shadow, meaning thereby the shadow constructed according to the laws of geometrical optics Owing to the fact that light consists of waves, the laws of geometrical optics are not strictly true, but the waves spread round the obstacle and encroach to some extent on the geometrical shadow. That they do not do so to a greater extent, was the principal difficulty of the wave theory in its earlier form This bending round of the waves has been called the "Diffraction" of light The simplest problems of Diffraction are those in which we imagine a plane or spherical wave to impinge on a plane perforated screen Whatever form or position the apertures HK, $H'K'$ (Fig 61) have, we can find the disturbance at a point P by Huygens' principle, if we know the disturbance at all points of the openings In the usual solutions of the problems, the assumption is made that the disturbance is the same at all points in the plane of the screen as it would be if the screen were away In other words, the screen simply obstructs the light which falls on its opaque portions, but does not otherwise alter the motion of the medium That the assumption is one which needs justification may be understood by contemplating *e g* the flow of water through a pipe, in which the stream lines are parallel straight lines, and imagining that at some place a diaphragm is introduced across the pipe, leaving only an aperture much narrower than its cross section We should here obviously arrive at erroneous results if we were to assume that the velocity of the water at all points of the opening has not been altered by the introduction of the diaphragm. In the case of the ordinary diffraction effects, it is found that the results arrived at by the simplified calculation are in agreement with experiment This is a consequence of the small size of the length of a wave of light as compared with the other linear magnitudes which enter into the calculation, the errors introduced being sensible only within a few wave-lengths of the obstacle.

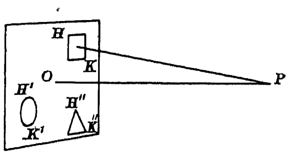

Fig 61.

We are allowed therefore to use Huygens' principle in its simple form, provided we correctly introduce the contribution which each small surface element s at a point S of the opening contributes to the amplitude at P. If r be the distance PS, ϕ the angle between r and

the perpendicular to the wave front at S, and θ the angle between r and the direction of vibration, the effect for homogeneous vibration of a small surface s at P is according to Stokes:

$$\frac{s(1+\cos\phi)\sin\theta}{2r\lambda}.$$

This expression is based on the assumption that the *displacements* in the openings are everywhere the same as if the screen were away. Lord Rayleigh, on the other hand, has shown that if the *forces* acting across the plane of the screen are the same as if the screen were absent, the effect of s would be

$$\frac{s\sin\theta}{\lambda r},$$

and has also pointed out that so far as the treatment of diffraction problems is concerned, the terms depending on θ and ϕ disappear in consequence of interference, so that we may with equal justice adopt the simpler expression arrived at in the previous article, and take the effect of an element at S to be according to convenience either $s/\lambda r$, or $s/\lambda p$, where p is the shortest distance from P to the wave front.

49. Babinet's principle. Two screens may be called complementary when the openings of one correspond exactly to the opaque portions of the other and vice versa. If b be the amplitude at P in the absence of any screen, and a_1, a_2 are vectors representing the vibration at P when either one or the other of two complementary screens is interposed, then the sum of the vectors a_1 and a_2 is obviously equal to b.

The principle due to Babinet allows us, whenever we have calculated the effect of one screen, to obtain that of the complementary screen without further trouble. A little care is necessary in using the principle, to take correct account of the difference in phase. But one simple result may at once be deduced from it. If a_1 is zero, a_2 must be equal to b. Hence at every point where there is no light with one of the screens, the intensity when the complementary screen is introduced, is equal to that observed when the light is unobstructed. This statement cannot however be reversed. If $a_1 = b$, a_2 may have any value between zero and $2a$. This is made obvious by the diagram (Fig. 62) in which OA represents the amplitude (b) of the unobstructed light, OB the equal amplitude (a) observed when one of the screens is introduced. BA is then that vector which together with OB has OA as resultant. If the point traces out the circle of radius a_1, the vector BA changes in magnitude from zero to $2a_1$.

Fig. 62

50. Shadows of a straight edge in parallel light Let a plane wave-front WF (Fig 63) fall upon a screen ME having a straight vertical edge passing through E, the plane of the drawing being horizontal, and let it be required to find the distribution of light on a distant and parallel screen SS' Draw the wave-front which passes through E, and divide up that portion EG of the wave-front which is not blocked out by the screen, into suitable zones, EP being the normal to the wave-front P lies on the edge of the geometrical shadow At P the active wave-front EG represents one of two exactly symmetrical halves of the complete wave-front, which would operate if the screen were away. Hence the introduction of the screen reduces the amplitude at the geometrical shadow to one half and the intensity to one quarter To find the amplitude at some point Q inside the geometrical shadow, construct Fresnel zones such that

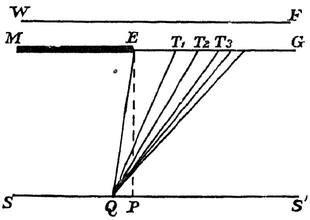

Fig 63

$$\frac{\lambda}{2} = T_1Q - EQ = T_2Q - T_1Q = T_3Q - T_2Q =$$

Unless Q is close to P, the resultant vibration due to the different zones will be alternately in opposite directions, and calling the effects of successive zones m_1, m_2, etc the total effect is

$$m_1 - m_2 + m_3 - m_4$$

In this case the values of m diminish too quickly to allow us to write down the sum as $\frac{1}{2}m_1$ It will however be some fraction of m_1, and as with increasing distances of Q from P, each of the zones diminishes in width, the effect at Q is the smaller the further that point lies inside the geometrical shadow The intensity which as has been shown is only 25 that of the incident light at the edge of the geometrical shadow, rapidly diminishes still further towards the inside of the shadow and soon becomes inappreciable

If the point Q lies outside the geometrical shadow the intensities are obtained by drawing the normal to the wave-front, and the Fresnel zones, according to Art. 47

The total effect in amplitude of that portion of the wave-front which lies to the right of the pole, when the shadow-throwing edge is on the left, is equal to 5, and the effect of the portion included between the pole and the edge is a maximum or a minimum, according as an odd or even number of zones are included between O and E (Fig 60) The first maximum takes place when Q is at such a distance from P that $OE = OT_1$. If the amplitude of the incident light is unity, and

the effects of successive zones are m_1, m_2, etc the first maximum has an amplitude $5 + m_1$, half the amplitude of the incident light being added to represent that complete part of the wave which lies to the right of O When Q has a position such that $OE = OT_2$, there is a minimum with an amplitude $5 + m_1 - m_2$. The next maximum has a value $5 + m_1 - m_2 + m_3$, and though the maxima and minima rapidly approach each other in magnitude the intensity continues to oscillate about its mean value as the point Q is moved away from the geometrical shadow The distances (x) of the maxima and minima from the edge are obtained from

$$x^2 = QE^2 - p^2 = \frac{4n-1}{4} p\lambda$$

The equation shows that the loci of the maxima and minima are parabolas.

TABLE IV

Shadow of straight edge.

Distance of screen $=100$, $\lambda = 5 \times 10^{-5}$, amplitude of incident light $=1$

No	Distance from edge in cms	Intensities: Outside geometrical shadow	Intensities: Inside geometrical shadow	m	$\sqrt{(4n-1)/2}$
1	061	1 3748	0298	1 217	1 225
2	094	7774	0140	1 873	1 871
3	117	1 1995	0091	2 345	2 345
4	137	8429	0067	2 739	2 739
5	154	1 1509	0053	3 082	3 082
6	170	8718	0044	3 391	3 391
7	184	1 1259	0037	3 674	3 674
8	197	8891	0033	3 937	3 937
9	209	1 1103	0029	4 183	4 183
10	221	9006	0026	4 416	4 416
11	232	1 0993	0024	4 637	4 637
12	·242	9092	0022	4 848	4 848
13	252	1 0910	0020	5 050	5 050

The angle x/p being proportional to $\sqrt{\lambda/p}$ is a small quantity unless n is large. But for large values of n the introduction of the screen causes no appreciable change in the distribution of light Hence the effect of the screen is confined to the neighbourhood of its geometrical shadow Table IV gives the intensities of light at the first seven

maxima and six minima outside the geometrical shadow, and the intensities inside at the same distances from the edge To give an idea of the scale, the positions to which the intensities refer are given for the case in which the shadow is received on a screen one metre away from the linear edge of the shadow-throwing object and the wave-length of light is 5×10^{-5} cms The meaning of the last two columns will be explained in Art 51

The table shows that at a distance of 2 5 mm from the edge of the geometrical shadow the light inside the shadow has only an intensity equal to the 500th part of that of the incident light, but that outside the shadow, at the same distance, the maximum and minimum intensities still differ by about 20%, while the interval between the bright and dark bands is 1 mm The light must of course be homogeneous if it is desired to see more than a few of the bands The distribution of the intensity of light in the neighbourhood of a straight edge is plotted in Fig 64 from the numbers given

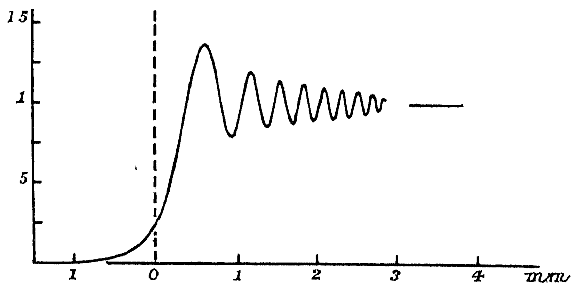

Fig 64

by Fresnel The dotted vertical line represents the edge of the geometrical shadow where the intensity is one quarter. The distance of the screen from the edge is one metre and the scale of abscissæ represents millimetres

51. Shadow of a straight edge in divergent light If L (Fig 65) represents the source of light, which we suppose to be a luminous line parallel to the edge E which throws the shadow, we may for simplicity take the beam to have a cylindrical wave-front with the luminous source as axis The traces of the wave-front with the plane of the paper are circles Drawing EG, the wave-front, passing through the edge, we may divide it into laminar Fresnel zones, OT_1, T_1T_2, etc which satisfy the condition that the

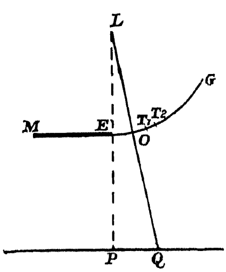

Fig 65

resultant of successive zones has opposite phases at Q. The distances of the edges of the laminae must be the same as in the previous article, so that

$$QT_n = QO + \frac{4n-1}{8}\lambda$$

The condition for the position of the maxima and minima is that a complete number of zones is exposed between O and E so that

$$QE - QO = \frac{4n-1}{8}\lambda$$

If we put $PQ = x$, $LE = LO = q$,

$$QE = \sqrt{p^2 + x^2} = p + \frac{x^2}{2p} \text{ app}$$

$$QO = LQ - q = p + \frac{x^2}{2(p+q)} \text{ app}$$

Hence the positions of the maxima and minima of light are determined by

$$\tfrac{1}{2}\frac{x^2 q}{p(p+q)} = \frac{4n-1}{8}\lambda,$$

which gives ·

$$x = \tfrac{1}{2}\sqrt{p\lambda(4n-1)(p+q)/q}$$

Fresnel in his celebrated Memoir on Diffraction obtained the expression

$$x = m\sqrt{p\lambda(p+q)/2q},$$

where m is a numerical factor which he calculated by means of the definite integrals which bear his name

To make our result agree with his, we must put

$$m = \sqrt{(4n-1)/2}$$

By means of his formula Fresnel obtained an excellent agreement between the observed and calculated positions of the maxima and minima, but the simple method which we have followed gives results which are sufficient for all practical purposes. To show that this is the case, the numerical values of the factor m calculated by Fresnel's method and ours respectively are entered into the two last columns of Table IV. All numbers except the first and second are identical, and even the difference in the position of the first band could hardly be detected by experiment.

As $LQ - QE$ is a constant for a given value of n, it follows that the loci of a maxima and minima are hyperbolas having L and E as foci. The width and hence the effect of each zone may easily be obtained and hence the intensities of the maxima and minima calculated, if desired

52. Shadow of a narrow lamina. If a cylindrical wave-front WF (Fig 66) falls on a vertical lamina of which AB represents the horizontal section, and throws a shadow on a screen MN, it is convenient to consider separately the portion of the screen HK which lies within the geometrical shadow and the two other portions which are respectively to the left and right of it Unless AB is very small, that portion of the wave which passes to the right of B does not affect very considerably the distribution of light to the left of H, and the distribution of light outside the geometrical shadow is therefore approximately that observed outside the shadow of a straight edge bordering a screen of unlimited extent To obtain the distribution of light at a point Q inside the geometrical shadow, construct the wave-front passing through A and B and divide it into Fresnel zones The resultant of the effects of all the zones to the right of B will agree in phase with that due to the first zone, and similarly for the light to the left of A the resultant phase must agree with that of the effect of the first zone There is a maximum or minimum of light at Q according as the phases resulting from the strips BT_1 and AR_1 act in conjunction or in opposition Unless Q is very near H or K the first zones may be drawn so that $QT_1 - QB = \frac{\lambda}{2}$ and $QR_1 - QA = \frac{\lambda}{2}$ In that case the first zones act in conjunction or in opposition according as $AQ - BQ$ is an even or odd multiple of half a wave-length The *positions* of the maxima or minima are therefore the same as if two dependent sources of light were placed at A and B The space HK is filled in consequence by equidistant bright and dark fringes, but except near the centre of the geometrical shadow the resultant amplitudes of the two portions of the active wave-front are not the same and there is therefore never complete darkness Near H and K the bands cease to be equidistant and gradually fuse into the ordinary fringes seen outside the shadow When the lamina is replaced by a thin wire or fibre, the distance between the internal fringes increases, and the position of the external fringes is no longer correctly calculated by considering only one portion of the wave-front. As the width of the obstacle is reduced, the fringes become less distinct and must disappear when the width is only a fraction of a wave-length, for in that case the obstruction is so small that the portions of the wave-front to the right and left of the obstacle cause an amplitude which must be practically identical with that of the unobstructed wave Plate II Fig 6 reproduces a photograph of the shadow of a wire and shows the central bright line

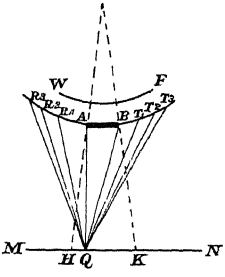

Fig 66

53. Passage of plane waves through a slit If a plane wave passes through a slit, placed parallel to the front of a wave, it is easy to obtain an expression for the distribution of light on a distant screen which is parallel to the first The edges of the slit being supposed vertical, let SS', Fig 67 and 68, be the intersection of the screen with a horizontal plane and subdivide the slit AB into a large number of vertical strips of equal width The illumination at a point P is equal to the sum of the effects of the separate strips If MM' be at a sufficient distance, all parts of the slit produce equal effects as regards magnitude, and the phase difference of the different rays is the same at the screen as on the arc of a circle AK drawn with P as centre For a distant screen this arc may be taken to be coincident with the line AK drawn at right angles to the direction of the rays (Fig 67) The phases of the rays proceeding from the centres of successive strips at the points where the rays cross the line AK are in arithmetic progression, and hence if the diagram of vibrations for the point P is constructed, we may apply the results of Art 5, so that if 2α be the phase difference between the vibrations due to the first and last ray, the resultant vibration has an amplitude $\frac{A \sin \alpha}{\alpha}$ where A is the amplitude at the central point To determine α, we require the phase difference corresponding to the optical distance BK, which if e be the width of the slit and θ the angle between the direction of the rays considered and the normal to the original wave-front is

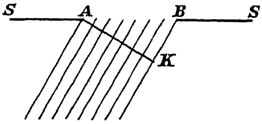

Fig 67

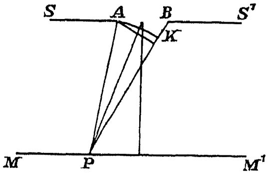

Fig 68

$$\alpha = \frac{\pi}{\lambda} e \sin \theta$$

The illumination at MM' is periodic, the amplitude being zero whenever α is a multiple of π, *i e* when $e \sin \theta$ is a multiple of λ To study the distribution of light more particularly, we must investigate the different values which the function $\left(\frac{\sin \alpha}{\alpha}\right)^2$ takes for different values of α Its zero values lie at equidistant intervals π The position of its maxima are found in the usual way from the condition

$$\frac{d}{d\alpha}\left(\frac{\sin \alpha}{\alpha}\right) = 0,$$

which gives
$$\alpha = \tan \alpha.$$

Draw the graph of $\tan\alpha$ as in Fig 69 The intersections of a straight line drawn at an angle of 45° to the coordinate axes, with the graph, determine the points for which $\tan\alpha = \alpha$ The figure shows that the points of intersection lie on successive branches of the graph and after the first lie near the positions for which α is an odd multiple of a right angle The first eight values of α for which $\sin\alpha/\alpha$ is a maximum are as follows

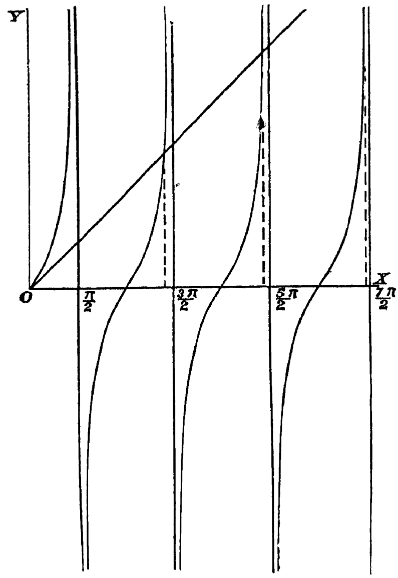

Fig. 69

$$\alpha_1 = 0,$$
$$\alpha_2 = 1·43\,\pi,$$
$$\alpha_3 = 2·46\,\pi,$$
$$\alpha_4 = 3·47\,\pi,$$
$$\alpha_5 = 4·48\,\pi,$$
$$\alpha_6 = 5·48\,\pi,$$
$$\alpha_7 = 6·48\,\pi,$$
$$\alpha_8 = 7·49\,\pi$$

The curve of amplitudes

$$A = A_0 \sin\alpha/\alpha$$

is drawn in Fig. 70 (dotted line) More important is the intensity curve $I = I_0 (\sin^2\alpha)/\alpha^2$ shown in the same figure Its coordinates,

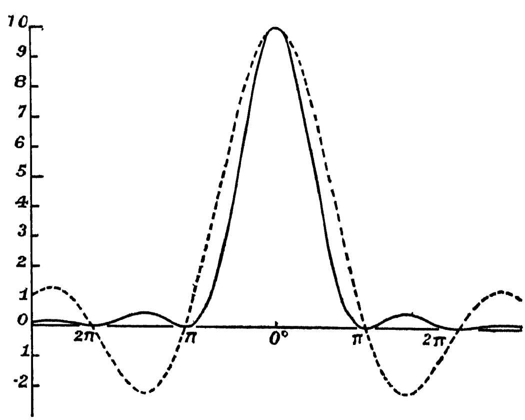

Fig 70

when I_0 is equal to one, are given in the third column of Table V. It appears that the bulk of the light is confined to values of α which

lie between $\pm\pi$, the intensity of the second maximum being less than $\frac{1}{20}$ of the intensity of that in the central direction. For the first minimum $(\alpha = \pi)$.

$$\sin\theta = \lambda/e$$

If e is equal to a wave-length, the light spreads out in all directions from the slit, with an intensity which is steadily diminishing as the inclination to the normal increases, but there are no other maxima of light beyond the central one. The equations must in that case be considered as approximate only, as is shown by the fact that the total intensity of light transmitted through the screen would according to the equations be less than the intensity of the light incident on the slit.

TABLE V.

α	$\sin\alpha/\alpha$	$(\sin^2\alpha)/\alpha^2$	α	$\sin\alpha/\alpha$	$(\sin^2\alpha)/\alpha^2$
0°	+1.0000	1.0000	270°	−0.2122	0.04503
15	+0.9886	0.9774	285	−0.1942	0.03771
30	+0.9549	0.9119	300	−0.1654	0.02736
45	+0.9003	0.8105	315	−0.1286	0.01654
60	+0.8270	0.6839	330	−0.0868	0.00754
75	+0.7379	0.5445	345	−0.0430	0.00185
90	+0.6366	0.4053	360	0.0000	0.00000
105	+0.5271	0.2778	375	+0.0395	0.00156
120	+0.4135	0.1710	390	+0.0735	0.00540
135	+0.3001	0.0901	405	+0.1000	0.01001
150	+0.1910	0.0365	420	+0.1181	0.01396
165	+0.0899	0.0081	435	+0.1272	0.01619
180	0.0000	0.0000	450	+0.1273	0.01621
195	−0.0760	0.00578	465	+0.1190	0.01416
210	−0.1364	0.01861	480	+0.1034	0.01069
225	−0.1801	0.03242	495	+0.0818	0.00670
240	−0.2067	0.04274	510	+0.0562	0.00315
255	−0.2170	0.04710	525	+0.0282	0.00080
			540	0.0000	0.00000

For values of e smaller than λ, the equations must not *a fortiori* be taken as giving more than an approximate representation of the facts, which may be wide of the truth if e is a small fraction of the wave-length.

When e is large compared with the wave-length, the whole light is confined to directions for which θ is very small. This explains the apparent discrepancy between the behaviour of sound and light, which retarded so long the general adoption of the undulatory theory of

light The amount of the spreading of waves which have passed through an opening depends entirely on the relation between the wave-length and the opening If sound-waves, having a length measured in feet, pass through an opening, the linear dimensions of which are of about the same magnitude, the waves expand in all directions, but if light-waves pass through the same openings, the spreading is practically nil, owing to the fact that the length of the waves is now very minute in comparison with the opening, and hence there is destruction of light by interference in oblique directions. To make experiments of sound and light waves comparable with each other, the openings should be made proportional to the lengths of the waves

54. Passage of light through slit. General case. In the previous article it has been assumed that the screen receiving the light is at a great distance We may now consider the more general case in which the screen is nearer and the incident light divergent. If Fig. 71 represents a horizontal section, L being the linear source and AB the aperture, we may find the amplitude at a point Q of the screen MM' by dividing the wave-front between A and B into appropriate zones Consider first the light at the central point P If O be the central point of the wave-front between A and B and the screen be at such a distance that $PA-PO=(4n-1)\lambda/8$ each half OA and OB of the wave-front contains an even or odd number of zones according as n is even or odd Hence there is a maximum or minimum of light at P according as n is odd or even As the screen is brought nearer, the observed system of fringes will alternately have a bright or dark centre at P. If p and q be the distances of P and L from the plane of the aperture, and d half the aperture of AB,

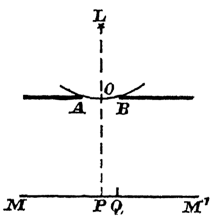

Fig 71

$$PO = p + q - \sqrt{q^2 + d^2}$$

$$= p - \frac{d^2}{2q} \text{ app.}$$

$$PA = \sqrt{p^2 + d^2}$$

$$= p + \frac{d^2}{2p} \text{ app}$$

$$\therefore \quad PA - PO = \frac{d^2}{2}\left\{\frac{1}{p} + \frac{1}{q}\right\}$$

and therefore

$$\frac{1}{p} + \frac{1}{q} = \frac{4n-1}{4}\frac{\lambda}{d^2}$$

determines the distance p of the screen from the opening, the central fringe being bright when n is odd and dark when n is even. When the point Q is not included in the geometrical beam of light which is bounded by the straight lines LB and LA, a similar reasoning leads to the conclusion that there is the centre of a bright or dark fringe at Q according as $AQ - BQ$ is an odd or even multiple of half a wave-length

55 Passage of light through a circular aperture When the perforations in a screen are such that we can divide the screen into circular zones, the calculation of the intensities is very simple for points in the axis of the zones

Let O (Fig 72) be the centre of a small circular aperture in a screen, and OP a line at right angles to the screen which we shall call the axis If it is required to determine the amplitude at P due to a wave-front of unit amplitude incident on the screen, which we shall consider in the first instance to be plane and parallel to it, we may divide the aperture into Fresnel's zones, which produce effects which are equal in magnitude but alternately opposite in direction If the radius OR of the aperture is such that an *even* number of zones is included, the amplitude at P is zero, if an *uneven* number is included the amplitude is a maximum and equal to that due to the first zone, and therefore double that of the unobstructed wave The introduction of the screen with small aperture doubles the amplitude therefore at certain points The condition for maximum or minimum of light is if $PO = p$, $OR = r$,

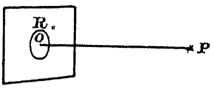

Fig 72

$$\frac{n\lambda}{2} = \sqrt{p^2 + r^2} - p = \frac{r^2}{2p} \text{ app}$$

where there is a maximum if n be odd and a minimum if n be even The general expression for the amplitude on the axis is found by subdividing the aperture into a large number of small zones of equal areas Their total effect, according to Art 5, is $(A \sin \alpha)/\alpha$ where for α we must put half the difference in phase at P of the disturbances due respectively to the first and last zone, i e half the difference in phase corresponding to an optical length $\frac{1}{2}n\lambda$ This gives

$$\alpha = \frac{\pi}{\lambda} \frac{n\lambda}{2} = \frac{\pi r^2}{2p\lambda}$$

A is the amplitude at P calculated on the supposition that the disturbances of all zones reach P in the same phase, which would according to Art 46 be $\pi r^2/p\lambda$, i e the area of the aperture divided by $p\lambda$.

The amplitude at P is therefore $2 \sin(\pi r^2/2p\lambda)$. The points of zero illumination which have already been determined are the nearer together the smaller the distance of P Sideways from the axis, the amplitudes cannot be calculated by simple methods, but general considerations similar to those which lead to accurate results in the case of long rectangular openings, are sufficient to show that there must be rhythmical alternations in the illumination Hence a screen placed across the axis will show bright and dark rings having at P a bright or dark centre according to the distance of P from the opening

The case of a divergent beam of light presents no further difficulty. We may subdivide the spherical wave-front into zones of equal area and obtain again at P the amplitude $\frac{A \sin \alpha}{\alpha}$ with the difference that $\alpha = \frac{\pi}{\lambda} r^2 \left(\frac{1}{2p} + \frac{1}{2q}\right)$, q being the distance of L from the screen. A has the same value as before Hence the points of maximum and minimum illumination are determined by

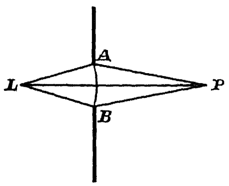

Fig. 73

$$\frac{1}{p} + \frac{1}{q} = \frac{n\lambda}{r^2},$$

and the amplitude at the maximum is $2q/(p+q)$

56. Shadow of a circular disc OR (Fig 74) being a circular disc, a spherical wave-front diverging from L, a luminous point on the axis of the disc, will throw a shadow on a screen SS', the centre of the shadow being on the axis If Fresnel zones are drawn on the wave-front, the total effect at P as regards amplitude may be determined as in Art 46 to be the same as that due to half the first zone, and if the disc is small, the first zone surrounding the edge of the disc has the same area as the central zone at O, which is covered by the disc Hence the illumination at P is the same as if the disc were away Round this central bright spot there are alternately dark and bright rings It will be an interesting exercise for the student to deduce the constancy of illumination on the axis of a shadow-throwing disc from Babinet's principle, making use of the amplitude at the bright and dark centres of the complementary circular aperture The fact that the shadow of a circular disc has a bright spot at its centre was discovered experimentally in the early part of the 18th century, but had been forgotten again when about 100 years later Poisson deduced it as a consequence of the wave-theory o

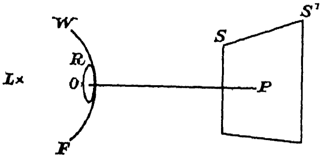

Fig 74

light Arago, who was unaware of the earlier experiment, tested Poisson's mathematical conclusion, and verified it

57. Zone plates. On a plane screen draw with O as centre, circles which divide the Fresnel zones with respect to a point P on the normal OP, the wave-front being supposed to be plane For the radii of the circle we have the relation

$$r^2 = np_0\lambda,$$

where p_0 is the distance OP, and where n takes the values 1, 2, 3 etc for successive circles Imagine the zones on the screen to be alternately opaque and transparent Then if a wave-front proceeding in the direction PO falls on the screen, the phases due to all transparent zones are in agreement at P, and hence the amplitude at P will be $\frac{1}{2}Nm$ where m represents the effect of the first zone and N the total number of zones

The amplitude at P will therefore be N times what it would be if the screen were away Such a zone plate acts like a lens concentrating parallel light to a focus, the focal distance being p_0 If now the source of light is moved to a point q from the screen, the zones will again unite their effects at P provided (Art 46)

$$\frac{1}{p} + \frac{1}{q} = \frac{n\lambda}{r^2},$$

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{p_0}$$

The relation between object and image is therefore the same as for a lens

Zone plates may be made by drawing circles on a sheet of paper, the radii of which are as the square roots of successive numbers, and painting the alternate zones in black When a photograph on glass is taken of such a drawing, a plate is produced which satisfies the conditions of a zone plate To prepare an effective zone plate involves great labour Prof R. W. Wood has published a reduced print of such a plate* from which other still more reduced copies may be prepared by photographic reproduction Prof Wood† has also described a photographic method by means of which zone plates may be made, which give for alternate zones a complete phase reversal A more perfect imitation of a lens may thus be obtained.

58. Historical. Augustin Jean Fresnel was born on May 10th, 1788, in Normandy, and entered the Government service as an engineer. He was occupied with the construction of roads, but lost his position owing to his having joined a body of men who opposed

* *Phil Mag.* XLV. p. 511. 1898 † *Ibidem*

Napoleon's re-entry into France, after his escape from Elba. Reinstated after Waterloo, he remained some time living in a small village in Normandy where his first study of the phenomena of diffraction seems to have been made. Fresnel was always of weak health and died on July 14, 1827. The undulatory theory of Optics owes to Fresnel more than to any other single man. His earlier work on Interference had to a great extent been anticipated by Thomas Young, but he is undoubtedly the discoverer of the true explanation of Diffraction. Young had tried to explain the external fringes of a shadow by means of interference of the rays which passed near the shadow-throwing object and those that were reflected from its surface. Fresnel, starting with the same idea, soon found that it was wrong, and proved by conclusive experiments that the surface reflexion had nothing to do with the appearance of the fringes. He then showed by mathematical calculation that the limitation of the beam, by the shadow-throwing object, was alone sufficient to cause the rhythmic variations of intensity outside the shadow.

CHAPTER VI.

DIFFRACTION GRATINGS

59. General theory of a grating. A grating is a surface having a periodical structure which impresses a periodical alteration of phase or intensity on a transmitted or reflected wave of light. The most common method of manufacturing a grating is to rule equidistant lines with a diamond point on a surface of glass or metal. The diamond introduces a periodical structure, each portion of which is probably very irregular, but which is repeated at perfectly regular intervals, Fig. 75. If the grating is ruled on a plane surface, that surface is called the plane of the grating. Any plane passing through corresponding points of the grooves such as A_1, A_2, A_3, is parallel to the plane of the grating. We distinguish between "reflexion gratings" and "transmission gratings" according as they are ruled on an opaque surface, the reflected or scattered light being used, or a transparent plate, through which the light is transmitted.

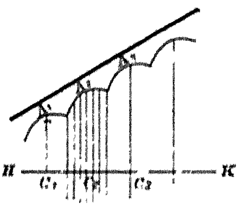

Fig. 75.

Let a plane wave front be incident parallel to the grating. Waves spread out from the different portions of the grooves which may be considered as centres of secondary disturbances. If the light be received on a distant screen, the resultant of all vibrations at each point may be determined. Consider that point of the screen which lies in a direction A_1C_1 from the grating, and draw a plane HK at right angles to that direction. As the optical distance from any point on HK to the corresponding point of the distant screen is the same, we may take the phases of the vibrations which are to be combined, to be the same as the phases at HK. We combine in the first place, those vibrations which are due to the secondary waves coming from one of the grooves. Selecting any point on the groove A_2, we may always express the phase of the resultant vibration due to the whole groove as that corresponding to an optical distance $A_2C_2 - \epsilon$,

where ϵ is some length which depends on the shape of the groove and on the direction of A_2C_2 The resultant amplitude similarly may be written ka, where a is the amplitude of the incident light and k a factor depending also on the shape of the grooves and the direction The different distances from the points of the groove to the plane HK do not affect the amplitude because that plane is only an auxiliary surface, the amplitude really being required at the screen which is so far away that the small differences in distance from different points of the grating are negligible Taking the resultant of the other grooves, we should find similarly that the resultant phases at HK may be derived from the optical distance $A_1C_1 - \epsilon$, $A_3C_3 - \epsilon$, etc, A_1, A_2, A_3, being corresponding points on the grating The theory of the grating depends on the fact that the values of ϵ and k are the same for each groove This involves the similarity of all the grooves, and if that similarity holds, the difference in phases between the resultant vibrations of two successive grooves is $(A_2C_2 - \epsilon) - (A_1C_1 - \epsilon)$ and is therefore independent of ϵ We may now draw a plane through any set of corresponding points of the groove and call it the plane of the grating (Fig 76), and in calculating the resultant phases at HK we need only consider the difference in the optical distance A_1C_1, A_2C_2, A_3C_3 If that difference is a multiple of a wave-length, the phases at HK are identical and we must then obviously have a maximum of light, wherever those identical phases are brought together This may either be the distant screen or the principal focus of a lens placed with its axis at right angles to HK The direction in which these maxima appear is easily obtained If θ be the angle between the normal to the grating and the direction A_1C_1 and A_1N be drawn at right angles to A_1C_1

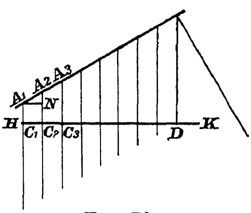

Fig 76

$$\sin\theta = \frac{A_2N}{A_1A_2} = \frac{n\lambda}{e} \qquad (1),$$

where e is the distance A_1A_2 between the grooves ruled on the grating, λ the wave-length and n an integer number The number of maxima is finite because $\sin\theta$ cannot be greater than one, and the highest value which we can take is therefore that integer which is nearest but smaller than e/λ If e were smaller than λ there could be no maximum except that for which $n = 0$ The amplitudes in the direction of the maxima are Nka, where N is the total number of grooves and k the constant already introduced, which may and does very seriously affect the amplitude It is theoretically possible that k is zero for one of the directions defined by (1) and in that case that maximum would of

course be absent It is also possible that k is unity, and in that case the whole of the light would be concentrated at or near that maximum.

The complete investigation of the grating includes the determination of the amplitudes of light in directions not necessarily confined to those at which the maxima appear We proceed, therefore, to find the distribution of light in the neighbourhood of the maxima The wave-length of a homogeneous beam incident on the grating being λ and having, as has been shown, a maximum in such directions that (Fig 76) $A_2N = n\lambda$, let the whole system of rays A_1C_1, A_2C_2 etc and with it the normal plane HK be turned round slightly so that A_2N now becomes $n\lambda'$, where λ' is a length differing little from λ. The difference in phase between the vibrations at C_2 and C_1 for the wave-length λ becomes $2\pi n\lambda'/\lambda$ or $2\pi n(\lambda' - \lambda)/\lambda$, as we may add or subtract any multiple of four right angles to a phase difference This is also the phase difference between the vibrations at C_3 and C_2, etc To obtain the complete resultant, we can therefore apply the proposition of Art 5, which gives for the amplitude of N vibrations of equal amplitude ka, and constant phase difference $2\alpha/N$, a resultant amplitude

$$Nka\frac{\sin\alpha}{\alpha}.$$

In the present case, $\alpha = \pi nN(\lambda' - \lambda)/\lambda$.

The distribution of intensity corresponding to this amplitude has been discussed in Art 53 Fig 70 shows for different values of α, the amplitude $(\sin\alpha)/\alpha$ (dotted curve) and the intensity $(\sin^2\alpha)/\alpha^2$ (full curve) The intensity has secondary maxima which are not, however, important compared with the principal one, at which $\alpha = 0$

The amount of light is everywhere small when α is greater than 2π, hence if Nn is large, the light is concentrated nearly in those directions for which $(\lambda' - \lambda)/\lambda$ is very small It is owing to the rapid falling off of the light at both sides of the principal maxima, that the grating can be made use of to separate the different components of non-homogeneous light, without any great overlapping of different wave-lengths

The condition for the first minimum $\alpha = \pi$, may be obtained in the most suitable form by considering that a series of waves with constant differences of path neutralize each other's effect, when the difference in optical length between the first and last is a wave-length There being N lines, the total difference in optical length is $Nn(\lambda' - \lambda)$, and for the first minimum this must be equal to λ' The condition that the minimum of light for a wave-length λ' is coincident with the maximum of a wave-length λ is therefore

$$\lambda'/(\lambda - \lambda') = nN \qquad \ldots \qquad (2).$$

It will be shown in Chapter VII that a spectroscope resolves a double line, the components of which have wave-lengths λ and λ', when the maximum of the diffraction image of one line coincides with the first minimum of the other The greater the value of Nn, the smaller is the difference $\lambda - \lambda'$ which may be resolved We may therefore take nN to be a measure of the resolving power

If we extend the above investigation to directions which are not near those of the maxima, the total light is found to be negligible, for a vibration diagram representing phases at C_1, C_2, by means of vectors OP_1, OP_2, and including all N vibrations, would have the points P_1, P_2, P_n distributed nearly symmetrically, so that the distance of their centre of gravity from the centre of the circle must be small compared with the radius of the circle

The incident wave-front has so far been taken as parallel to the plane of the grating For oblique incidence, consider a grating formed by ruling lines on a glass surface, and let a plane wave be transmitted obliquely through it Let A_1, A_2 (Fig 77) be corresponding points on successive grooves, and LM the incident wave-front, inclined at an angle ϕ to the plane of the grating Draw two rays LA_1, MA_2, and consider the light diffracted in the direction A_1C_1, inclined at an angle θ to the normal of the grating Draw A_1N and A_2T at right angles to A_1C_1 and A_1L respectively The difference in phase between C_1 and C_2 is then

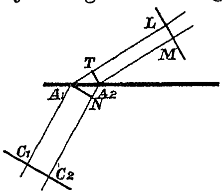

Fig 77

$$e(\sin\phi - \sin\theta),$$

and there is a maximum when

$$e(\sin\phi - \sin\theta) = \pm n\lambda \qquad (3).$$

θ and ϕ are here taken as having the same sign when they are both on opposite sides of the normal

Writing γ for $\phi - \theta$, the angle between the incident and diffracted beams, the condition for a minimum or maximum of deviation is $\frac{d\gamma}{d\theta} = 0$, which leads to $d\phi = d\theta$. By differentiating (3) we obtain

$$\cos\phi d\phi - \cos\theta d\theta = 0$$

If $d\phi = d\theta$ it follows that $\cos\phi = \cos\theta$, *i e* $\phi = \pm\theta$. ϕ and θ cannot be equal unless $n = 0$, which case need not be considered. For the condition of maximum-minimum we have therefore $\phi = -\theta$, which shows that the incident and diffracted light form equal angles with the plane of the grating. Further consideration shows that it is a minimum and not a maximum deviation that is involved

If $\phi = -\theta$ the deviation is 2θ Equation (3) becomes in that case

$$2e \sin \frac{\gamma}{2} = 2e \sin \theta = n\lambda$$

60. Overlapping of spectra. The maxima of light for normal incidence have been shown to take place when $e \sin \theta = n\lambda$ For each value of n, the maxima of the different wave-lengths take place along different directions, and hence the grating "analyses" the light which falls on it and produces homogeneous light It acts in this respect like a prism, but splits up the light into a number of spectra, each value of n giving a separate spectrum For $n = 0$, there is a maximum, but there is no spectrum because the position of the maximum is independent of the wave-length The direction of this maximum is the direction of the incident light in a transmission grating, or in a grating which acts by reflexion, it is the direction in which the incident beam would be reflected from a polished surface coincident with the grating For $n = 1$, we have the so-called spectrum of the first order, which spreads over the quadrant between $\theta = 0$ for $\lambda = 0$ and $\theta = \frac{1}{2}\pi$ for $\lambda = e$ Similarly the spectrum of the second order, for which $n = 2$, spreads over the same quadrant, the limits of wave-length being $\lambda = 0$ for $\theta = 0$, and $\lambda = e/2$ for $\theta = \frac{1}{2}\pi$ For each value of θ we have therefore an infinite number of overlapping maxima corresponding to all wave-lengths which are submultiples of $e \sin \theta$ If we confine ourselves to eye-observations, we need only consider the wave-lengths lying between 4×10^{-5} and 8×10^{-5}. The limits θ' and θ of the spectra of different orders are then

$$\begin{aligned}
&\text{for } n = 1, \quad 4 \times 10^{-5} = e \sin \theta' \text{ and } 8 \times 10^{-5} = e \sin \theta, \\
&\text{for } n = 2, \quad 8 \times 10^{-5} = e \sin \theta' \text{ and } 16 \times 10^{-5} = e \sin \theta, \\
&\text{for } n = 3, \quad 12 \times 10^{-5} = e \sin \theta' \text{ and } 24 \times 10^{-5} = e \sin \theta, \\
&\text{for } n = 4, \quad 16 \times 10^{-5} = e \sin \theta' \text{ and } 32 \times 10^{-5} = e \sin \theta.
\end{aligned}$$

In Fig 78 the extension of the different spectra is marked by straight lines lying above each other to avoid actual overlapping If the wave-lengths marked are those corresponding to the first order spectrum, we may obtain the wave-length of the spectrum of order n, by dividing these numbers by n

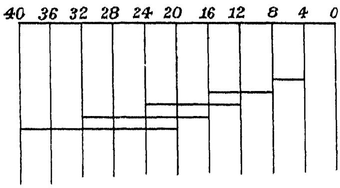

Fig 78

The visible spectrum of the first order stands out clear of the rest, but the second and third overlap to a great extent, the range between $\lambda = 6 \times 10^{-5}$ and $\lambda = 8 \times 10^{-5}$ of the second order being coincident with the range of $\lambda = 4 \times 10^{-5}$ to $\lambda = 5.3 \times 10^{-5}$ of the third order The spectra of higher orders spreading over greater ranges of θ overlap

more and more, and special devices have to be adopted to separate the spectra, when observations are made in the higher orders When spectra are to be recorded by photography, there is a similar overlapping but its range is different

61. Dispersion of gratings. The maxima of two wave-lengths λ_1 and λ_2 being in such positions that

$$e \sin \theta_1 = n\lambda_1,$$
$$e \sin \theta_2 = n\lambda_2,$$

the ratio $(\theta_1 - \theta_2)/(\lambda_1 - \lambda_2)$ may be taken to measure the angular dispersion of the grating The ratio increases with increasing values of $n\lambda$ and hence the dispersion increases with the order of the spectrum.

If the incident beam is oblique

$$e(\sin \theta - \sin \phi) = n\lambda,$$

which, by differentiation, gives with a constant value of ϕ

$$e \cos \theta \, d\theta = n d\lambda,$$

so that in this case the angular dispersion is $\frac{d\theta}{d\lambda} = n/e \cos \theta$.

When the diffracted beam leaves the grating nearly normally, $\cos \theta$ varies much less rapidly than $\sin \theta$ In that case the dispersion is proportional to the order of the spectrum and independent of the wave-length, *i e* equal angular separation means equal differences of wave-length We then say that the spectrum formed is "normal"

62. Resolving power of gratings. The use of a grating as an analyser of light depends on its power to form a pure spectrum. To obtain a measure of the purity of a spectrum, we may imagine it to be projected on a screen, which has a narrow opening parallel to the original slit intended to transmit only that wave-length which has a maximum coinciding in position with the opening It is then found

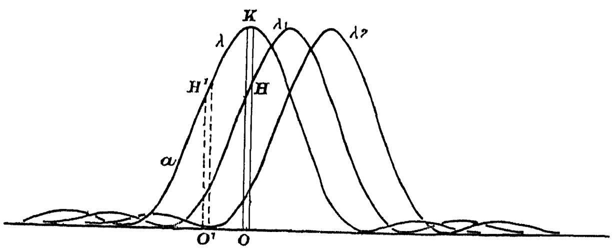

Fig 79

that the waves passing through even an indefinitely narrow aperture are not absolutely homogeneous In Fig 79 the curve a represents the

distribution of light on the screen for a given wave-length OK indicates the position of a narrow opening placed so as to transmit the maximum amount of light having a given wave-length λ, the amount so transmitted being proportional to the intensity OK and to the width of the opening If λ_1 be a wave-length near λ, it will have its maximum a little to one side Its intensity curve is represented by the second curve and an amount of its light proportional to OH passes through the opening The curves of intensity having no definite limit, there is some light of every wave-length passing through the slit, but the intensity quickly diminishes and we need only consider those wave-lengths which are not very different from λ If we wish to compare different spectrum-forming instruments with each other, it will be sufficient to limit the investigation to that light which lies between the two minima on either side of the maximum

It follows from Art 59 that a wave-length λ_1 has its first minimum when there is maximum for λ if $nN(\lambda_1 - \lambda)/\lambda = \pm 1$ Hence we may say that the range of wave-lengths passing through the opening extend from a wave-length $\lambda\left(1 + \frac{1}{Nn}\right)$ to a wave-length $\lambda\left(1 - \frac{1}{Nn}\right)$ The quantity Nn has been called the resolving power of the grating Denoting it by R, we may say that very little light passes through the slit which differs in wave-length from λ by more than λ/R Resolving power will be further considered in Chapter VII.

63. Wire gratings In certain cases, the intensity of the spectra of different orders may be calculated. If the grating is formed by a number of equidistant thin wires of equal thickness (Fig 80), the periodicity of the grating is such that one portion does not obstruct the passage of the light whilst the other is opaque Take the incident light to be normal to the grating, and let the widths of each transparent and opaque portion be a and b respectively, the amplitude of the light diffracted at an angle θ to the normal is then (Art 53) $(A \sin \alpha)/\alpha$ where $\alpha = \pi a \sin \theta/\lambda$.

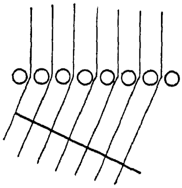

Fig 80

The maximum of the nth order is determined by

$$(a + b) \sin \theta = n\lambda \text{ ; so that } \alpha = \pi a n/(a + b).$$

The amplitudes at the maxima are therefore

$$\frac{A\,(a + b)}{\pi n a} \sin \frac{n\pi a}{a + b}$$

For the central image, in which there is no dispersion, $\alpha = 0$ and the amplitude is A The law of falling off in the intensities of the different images to the sides of the maxima is the same near all maxima,

so that for the ratio of the intensities of the images, we may substitute the ratio of the squares of the amplitudes at the maxima For the calculation of the amplitude at the central maximum, it is sufficient to point out that the interposition of the grating reduces the amplitude in the ratio of its transparent portion to its total surface, *i e* in the ratio $a/(a+b)$, and hence the intensity of the central image is $\{a/(a+b)\}^2$, if the intensity of the incident light is unity. This determines the value of A

We now obtain for the intensities of the other images,

$$\frac{1}{n^2\pi^2}\sin^2\left(\frac{n\pi a}{a+b}\right)$$

If $a=b$, the sine factor is zero for all even values of n, so that the spectra of even order disappear, and the intensities of the spectra of odd orders are, in terms of the incident light, $\frac{1}{\pi^2}, \frac{1}{3^2\pi^2} \quad \frac{1}{n^2\pi^2}$

The fraction $1/\pi^2$ represents the maximum intensity which the spectrum of the first order can possibly have in this class of gratings, and shows what a considerable amount of light is lost when a grating is used as an analyser of light If we desire to make the second order spectrum as intense as possible, we must make a/b equal to 1/3 or 3, but even in this case, we should only secure little more than two per cent of the light

It is instructive to note that the grating reduces the *intensity* of the total light transmitted in the ratio $a/(a+b)$, which is also the ratio in which the *amplitude* of the central image is reduced The difference between $a/(a+b)$ and $\{a/(a+b)\}^2$ gives the amount of light which goes to form the lateral spectra

64. Gratings with predominant spectra. Rulings of gratings may be devised which concentrate most of the light into one spectrum. Fig 81 represents the section of such a grating ruled on a reflecting surface. If the oblique portions of the grating are at such an angle that light incident in the direction of the arrow would, by the laws of geometrical optics, be reflected in the direction A_1C_1 then all the rays from each of the oblique portions would be in equal phase at a plane HK, drawn at right angles to A_1C_1 If, further, the difference in optical length at HK between A_2C_2 and A_1C_1 be a wave-length, there is coincidence of phase between the rays from successive rulings.

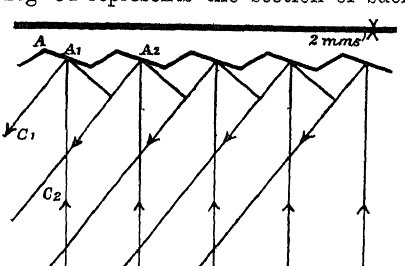

Fig 81

Hence the amplitude at a point which is at the same optical distance from HK (*e.g.* the focus of a lens adjusted for infinity) is the same as if the whole wave-front HK were reflected in the ordinary way. The resultant amplitude is therefore less than the resultant amplitude of the incident wave, only on account of the contraction in the width of the beam due to obliquity. If θ be the angle between A_1C_1 and the incident beam, it would follow that the intensity of the first order spectrum is $\cos^2\theta$ in terms of the resultants of the incident light. This loss of light is accounted for by the light reflected from the other set of inclined faces. If the ruling is such that the first order spectrum is at an angle of 30° from the normal, three-quarters of the whole light would go to form that spectrum. For normal incidence we have as before, $\sin\theta = \lambda/e$, and the reflecting facets must be inclined at an angle $\theta/2$. The condition for maximum light can only be fulfilled for one wave-length at a time, but a slight tilting of the grating supplies the means of adjustment for any desired wave-length. Transmission gratings may be ruled on the same principle, the condition being that the angles of the inclined facets are such that the incident rays in each little prism formed are refracted along paths at right angles to HK, and that there is a retardation of a wave-length between two corresponding rays A_2C_2 and A_1C_1. Mr T. Thorp has been able to demonstrate the practical possibility of manufacturing gratings of the kind considered. Triangular grooves were cut in a metallic surface, and a layer of liquefied celluloid was allowed to float and solidify over this grooved surface. On removal, the celluloid film showed in transmitted light spectra which were all very weak except that of the first order on one side. HK (Fig. 82) gives the direction of the wave-front of the diffracted wave which carries the maximum intensity for the wave-length λ.

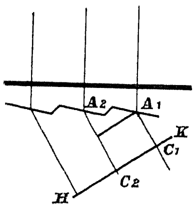

Fig. 82

65. Echelon gratings. If a reflecting grating were constructed on a principle similar to that of the last article, but subject to the additional condition that rays which go to form a particular spectrum return along the path of the incident light, the spectrum formed by reflexion would contain the whole intensity of the incident light. This consideration leads to Michelson's echelon grating. In Fig. 83 let a number of plates, T_1, T_2, T_3, etc. be placed so that the different portions of a wave-front WF are reflected back parallel

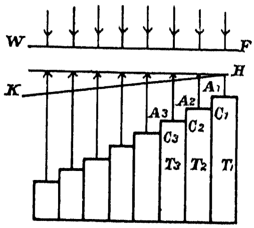

Fig. 83

to themselves from each of the plates, then if the depths of the steps A_1C_1, A_2C_2, A_3C_3, are all equal to $n\lambda$, a multiple of a wave-length, the reflected beam has intensity equal to the incident beam, neglecting the loss of light at reflexion For that particular wave-length, there cannot therefore be light in any other direction The reasoning holds for all those wave-lengths for which the step is an exact multiple of a wave-length, and we may, if n is great, have a great number of maxima of light all overlapping in the same direction

At a surface HK inclined to WF at a small angle θ, the retardation of successive corresponding rays is $e\theta$, where e is the width of each step Hence there is coincidence of phase for a wave-length λ' at corresponding points of HK if

$$n(\lambda - \lambda') = e\theta$$

For the dispersion $\theta/(\lambda - \lambda')$ we thus obtain n/e But only a very small part of each spectrum is visible because the intensity of light falls off very rapidly to both sides of the normal direction

At a wave-front parallel to WF, the relative retardation of two waves λ and λ', for the light reflected by the last element, is $Nn(\lambda - \lambda')$ if there is coincidence of phase for light reflected at the first element. Hence equation (2) holds, and the resolving power is Nn, as with ordinary gratings

A reflecting grating of the kind described would be difficult to construct, but excellent results have been obtained by Michelson with a transmission grating based on the same principle

A number of equal plates of thickness t are arranged as in Fig 84. Each part of the beam is retarded by $(\mu-1)t$ more or less than its neighbour For normal transmission, there is equality of phase everywhere on WF if

$$(\mu - 1)t = n\lambda$$

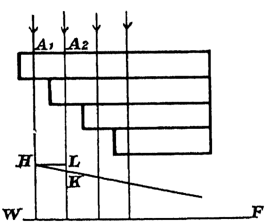

Fig 84

If for the wave-length λ', having refractive index μ', the diffracted wave-front is parallel to HK, the phase at H and K must be the same, or

$$(\mu' - 1)t + LK = n\lambda',$$

and if e be the distance between corresponding points A_1, A_2, the angle θ through which the front is turned is LK/e or

$$\theta = \{n\lambda' - (\mu' - 1)t\}/e$$
$$= \{n(\lambda' - \lambda) - (\mu' - \mu)t\}/e$$

The angular dispersion is therefore

$$\theta/(\lambda' - \lambda) = \left\{n - \left(\frac{\mu' - \mu}{\lambda' - \lambda}\right)t\right\}\Big/e \quad . \quad (4)$$

If N be the total number of plates, the first minimum of the diffractive image of λ' coincides with the maximum of λ, if the total retardation $Ne\theta$ is equal to λ Hence multiplying both sides of (4) by Ne, we find

$$\frac{\lambda}{\lambda' - \lambda} = N\left(n - \frac{d\mu}{d\lambda} t\right) \qquad (5),$$

where $\frac{d\mu}{d\lambda}$ has been substituted for $(\mu' - \mu)/(\lambda' - \lambda)$, as only very small variations of μ and λ come into play

Substituting $n\lambda = (\mu - 1) t$, the ratio of the second term on the right-hand side of (5) to the first is $\frac{\lambda d\mu}{d\lambda}\Big/(\mu - 1)$, and this for flint glass, and in the centre of the visible spectrum varies between about -05 and -1 We may therefore say that the value of $\lambda/(\lambda' - \lambda)$ for this form of grating is from 5 to 10 per cent greater than Nn, but approximately the resolving power is the same as for the nth order of an ordinary grating having a total number N of grooves Full intensity is only obtained for those wave-lengths for which $t = n\lambda$ But a slight tilting of the grating increases the effective thickness t, and brings any desired wave-length into the best position The total light is, however, in any case, confined to the immediate neighbourhood of the direction of the incident light, because the width of each element is large compared with a wave-length It is worth while to discuss this a little more closely The angular distance between the principal maximum and the first minimum with an aperture e is according to Art 62, λ/e We may therefore, disregarding the light which is beyond the first minimum, say that the spectra have appreciable brightness only to a distance λ/e on the two sides of the normal. Consider now that the maximum of the nth order of λ' coincides with the maximum of the $(n+m)$th order of λ when $n\lambda' = (n + m)\lambda$ If in (4) we neglect the second term on the right-hand side and for θ substitute $2\lambda/e$ which measures the total angular space within which the light has an appreciable intensity we find $2\lambda = (\lambda' - \lambda) n$ or $\lambda' = (n + 2)\lambda$, which by comparison shows that $m = 2$. No order except n, $n + 1$ and $n + 2$ can therefore be visible In the case considered the orders n and $n + 2$ would just coincide in position with the places of zero illumination and the central image would contain all the light As a rule there will be two spectra As regards intensity of light, the echelon form gets rid of one of the chief difficulties in the use of gratings, as the light must be concentrated almost entirely into two spectra, and we may adjust the grating so that the intensity is practically confined to one spectrum only

The overlapping of spectra of different orders is, however, a serious inconvenience, for it must be remembered that although for each wave-

length there are only two orders visible, the number of the order is different for the different wave-lengths, and the total number of overlapping orders is very great As an example, consider normal incidence on a grating having its plates of thickness 5 cm For a wave-length $\lambda = 5 \times 10^{-5}$, the thickness is 10,000 times the wave-length, so that we should observe a spectrum of the 10,000th order. Coincident with it, and for a slightly differing wave-length, we should have the spectra of orders which are near that number Thus $n\lambda = 5$ is satisfied for $n = 8{,}000$, if $\lambda = 6\,25 \times 10^{-5}$ There are therefore 2000 coincident maxima within the range of wave-lengths 5×10^{-5} and $6\,25 \times 10^{-5}$, the former lying in the green and the latter in the orange

These overlapping spectra must be separated or got rid of This is done by means of an ordinary spectroscope, which can be used in two ways In the form of the apparatus as it is most commonly constructed, the light is sent through a train of prisms before it falls on the slit of the echelon collimator The resolving power of the prisms should be sufficient to exclude all light belonging to the maxima which it is desired to exclude We may also use a train of prisms to separate the maxima *after* they have passed through the echelon, and this arrangement, which would seem to possess some advantages, was apparently used by Michelson in his first experiments

66 Concave gratings That certain gratings possessed a focussing power had been noticed by a number of observers, and the explanation of the fact presents no difficulties, but what previously had always been considered a defect to be avoided, became in the mind of Rowland an object to be desired, and by very perfect mechanical contrivances was made use of to revolutionize spectroscopic research

It is always possible to construct theoretically the ruling of gratings on surfaces of any shape, such that an image of a spectrum at any desired point shall be formed

Let A (Fig 85) represent a point source of light, and let it be desired to form an image of the spectrum of the first order so that all the light of wave-length λ shall be concentrated at B With A and B as foci, draw ellipsoids such that if P, P', P'' be points on successive ellipsoids,

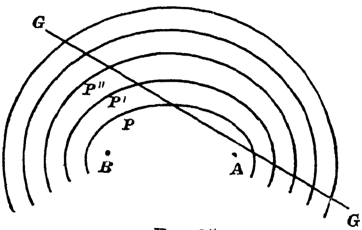

Fig 85

$$AP + PB = m\lambda,$$
$$AP' + P'B = (m + \tfrac{1}{2})\,\lambda\,,$$
$$AP'' + P''B = (m + 1)\,\lambda, \text{ etc}$$

Let GG' be the trace of the surface which it is desired to convert into a

grating The grating intersects these ellipsoids in curves which divide it into Fresnel zones The light which might reach B from successive zones is in opposition and no luminous disturbance can therefore exist at that point But if some change be made in the zones, so that the amount of light scattered by alternate zones is either obliterated or at any rate weakened, the plate will act like a zone plate and light will be focussed at B Ruling lines with diamond point parallel to the lines of division between the zones and at distances equal to the distance between alternate zones, is sufficient to produce the desired effect As the construction of the zones depends on the wave-length, the spectrum formed has a focus at B for a particular wave-length only But the adjoining wave-lengths are concentrated into other foci in the neighbourhood If we desire to produce spectra of higher orders, we may draw the zones so that the sum of the distances of any point from A and B is $m\lambda$, $(m+n)\lambda$, $(m+2n)\lambda$, etc. If a portion of the space filled by each zone so formed is cut by a diamond, so that the corresponding portions of all zones are modified in like manner, a source of light at A produces a spectrum of the nth order at B

In practice, we are confined to rulings in straight lines on plane or spherical surfaces We are also unable to rule the lines *accurately* except by means of a screw turned step by step through equal angles. It is Rowland's discovery that gratings with very small aberrations can be made by ruling lines on a spherical surface by means of a screw. In Fig 86 let A represent a source of light, and B the point at which it is desired to form a spectrum of the nth order We confine the investigation to rays lying in the plane containing AB and the normal OC of a curved grating GG', C being the centre of curvature Take OC as axis of X, and the tangent to the grating at O as axis of Y

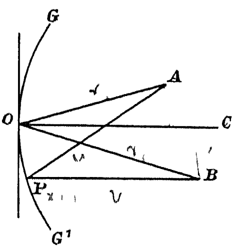

Fig 86.

Put $OA = r$, $BO = r_1$, $AP = u$, $BP = v$

If P lies on the edge of the mth zone, and if the nth order spectrum is in focus at B,

$$u + v = r + r_1 \pm mn\lambda.$$

If the distance between successive rulings is such that its projection on OY is constant and equal to e, $y = me$, hence eliminating m,

$$u + v = (r + r_1) \pm \frac{n\lambda y}{e} \qquad (6)$$

If this condition could be fulfilled absolutely we should have a perfect image at B It must be our object now to see how nearly we may satisfy equation (6) in practice.

Writing a, b, for the coordinates of A, a_1, b_1, for those of B, we have

$$u^2 = (y-b)^2 + (x-a)^2$$
$$= r^2 + x^2 + y^2 - 2by - 2ax \qquad (7)$$

If ρ is the radius of curvature of the grating, the equation to the circle in which the grating cuts the plane of the paper gives

$$2\rho x = x^2 + y^2$$

But our investigation may be made to include gratings, deviating from the spherical shape, so long as the osculatory circle at O has a radius of curvature ρ We therefore more generally put the equation of the trace of the grating

$$2\rho x = \beta x^2 + y^2 \qquad (8),$$

where β is a numerical constant which is *one* in the case of a sphere. Combining (7) and (8) we obtain by simple transformations

$$u^2 = \left(r - \frac{by}{r}\right)^2 + \left(\frac{a}{r^2} - \frac{1}{\rho}\right) ay^2 + \left(1 - \frac{a\beta}{\rho}\right) x^2 \qquad (9)$$

The second term is of the second order of magnitude as regards y, and the third term of the *fourth* order of magnitude as regards the same quantity Retaining only quantities of the second order,

$$u = \left(r - \frac{by}{r}\right) + \frac{1}{2r}\left(\frac{a}{r^2} - \frac{1}{\rho}\right) ay^2$$

Similarly
$$v = \left(r_1 - \frac{b_1 y}{r_1}\right) + \frac{1}{2r_1}\left(\frac{a_1}{r_1^2} - \frac{1}{\rho}\right) a_1 y^2$$

In order that the grating should fulfil its object, it is necessary that at least to this order of magnitude, (6) should be fulfilled. Hence substituting u and v into that equation and putting the factors of y and y^2 equal to zero we obtain

$$\frac{b}{r} + \frac{b_1}{r_1} = \mp \frac{n\lambda}{e} \ldots \qquad (10),$$

and
$$\frac{a}{r}\left(\frac{a}{r^2} - \frac{1}{\rho}\right) + \frac{a_1}{r_1}\left(\frac{a_1}{r_1^2} - \frac{1}{\rho}\right) = 0 \qquad (11)$$

The first condition defines the direction in which the diffracted image lies, for if ϕ and θ are the angles which AO and BO make respectively with the normal, $r \sin \phi = b$, and $r \sin \theta = b_1$, and (10) is therefore identical with

$$e(\sin \theta + \sin \phi) = \pm n\lambda$$

This equation is therefore common to the curved and plane grating The second condition now gives the *distance* of the diffracted image, for as $r \cos \phi = a$, $r \cos \theta = a_1$, (11) is identical with

$$\frac{\cos^2 \phi}{r} + \frac{\cos^2 \theta}{r_1} = (\cos \phi + \cos \theta) \frac{1}{\rho}$$

If θ and ϕ be equal and small, this is the well-known relation between object and image of a concave mirror

We must now try to see to what order of magnitude we can get rid of aberrations Leaving out terms of the fourth order equation (9) may be written

$$u^2 = \left(r - \frac{by}{r}\right)^2 + \left(\frac{a}{r^2} - \frac{1}{\rho}\right) ay^2,$$

and hence

$$u = \left(r - \frac{by}{r}\right) + \tfrac{1}{2}\frac{\left(\frac{a}{r^2} - \frac{1}{\rho}\right)}{r - \frac{by}{r}} y^2 + \text{terms of higher orders}$$

The term containing y^3 disappears if

$$a\rho = r^2,$$

and as (11) must be satisfied, this involves also

$$a_1 \rho = r_1^2$$

The first equation places the source of light on a circle of radius 2ρ with its centre in the line OC, and the second equation shows that the same circle contains the image B

Limiting ourselves to this circle for the position of the source, (9) becomes

$$u^2 = \left(r - \frac{by}{r}\right)^2 + \left(1 - \frac{\beta a}{\rho}\right) x^2,$$

and to quantities of the fourth order,

$$u = r - \frac{by}{r} + \frac{x^2}{2r}\left(1 - \frac{a\beta}{\rho}\right),$$

$$v = r_1 - \frac{b_1 y}{r_1} + \frac{x^2}{2r_1}\left(1 - \frac{a_1\beta}{\rho}\right)$$

Comparison with (6) shows that the terms of the fourth order depend on the factor

$$\frac{1}{r}\left(1 - \frac{a\beta}{\rho}\right) + \frac{1}{r_1}\left(1 - \frac{a_1\beta}{\rho}\right) \quad \ldots (12).$$

In the position in which Rowland's gratings are generally used $a_1 = \rho = r_1$ and $a = \rho\cos^2\phi = r\cos\phi$ Hence (12) reduces to

$$(1 + \cos\phi)(\sec\phi - \beta)/\rho \quad . \quad (13)$$

The terms of the fourth order cannot be got rid of therefore except for a particular value of ϕ For spherical gratings $\beta = 1$ and the second factor of (13) is small for small values of ϕ, so that the aberration is least important for the spectra of lower orders It could be corrected entirely for a particular value of ϕ by making $\beta = \sec\phi$, but this would involve departure from a spherical surface

The outstanding error of optical length for spherical gratings is obtained by restoring in (12) the dropped factor $x^2/2$ or $y^4/8\rho^2$. The error then reduces with $\beta=1$ to

$$\frac{x^2}{2\rho}\sin\phi\tan\phi \quad \text{or} \quad \frac{y^4}{8\rho^3}\sin\phi\tan\phi$$

The maximum error of optical length may be as much as a quarter of a wave-length without seriously damaging the definition Hence if y is half the width of the spectrum, we have for the condition of still perfect definition,

$$\frac{y^4}{2\rho^3}\sin\phi\tan\phi<\lambda,$$

and if $\sin\theta=0$ and $\sin\phi=\dfrac{n\lambda}{e}$, the greatest value of y which is half the width of the grating should not exceed

$$\left(\frac{2\rho^3 e\cot\phi}{n}\right)^{\frac{1}{4}}=\rho\left(\frac{2e\cot\phi}{n\rho}\right)^{\frac{1}{4}}$$

The dispersion (Art 61) of the grating is $n/e\cos\theta$, and for $\theta=0$ is therefore independent of the wave-length The grating used in such a way that the spectrum appears on its axis forms, therefore, a normal spectrum.

Rowland's method of mounting the grating, which combines the advantages of maximum definition and the formation of a normal spectrum, is shown diagrammatically in Fig 87 G is the grating and is held by a rigid beam GC of length equal to the radius of curvature of the grating, which carries at its other end the photographic camera HK AS and BS are two strong beams placed at right angles to each other and carrying rails which support two carriages which can roll along the beams and support in their turn the beam GC which is pivoted on them The slit is placed at S As G, S, C lie on a circle of diameter ρ, a luminous source at S will always have its image at C, when the proper position of the beam GC has been found As the beam is rolled along the rails, successive wave-lengths and successive spectra make their appearance in proper focus at C For a given position of the beam, the focus for the different wave-lengths lies on a circle of diameter ρ, and the photographic plate HK must therefore be bent to a curvature equal to twice that of the grating The angle CGS is the angle called ϕ above,

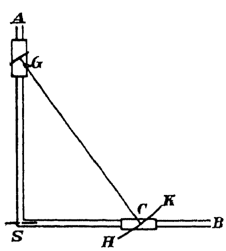

Fig 87

and as $\theta=0$, $\qquad e\sin\phi=n\lambda,$

and as $\qquad SC=\rho\sin\phi,$

it follows that $\qquad SC=\dfrac{n\lambda\rho}{e}.$

The beam SB may therefore be divided into a scale of wave-lengths by equal divisions, and the wave-length which occupies the centre of the field at C may be read off directly on that scale

A complete discussion of the theory of the concave grating which we have in great part followed here, has been given by Runge, and published by Kayser (*Spectroscopie*, Vol I. p 400) The same volume contains valuable information on the methods of adjustment and on the literature of the subject. It should however be mentioned that though later investigations have simplified the analysis, the essential points of the theory are all contained in Rowland's* original papers

67 Measurement of wave-length Plane gratings allow us to measure very accurately the length of a wave of light (λ) In Fig 93 C represents a collimator which admits the light through a narrow slit

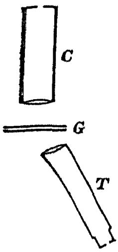

Fig 88

The source must be one sending out nearly homogeneous radiations, such as a tube filled with a vapour under reduced pressure rendered incandescent by the electric discharge The light is allowed to fall on a grating at G (Fig 88) and is observed by means of a telescope T If the axis of the telescope coincides with the direction of a maximum of light in the diffracted beam, we have the relation

$$e(\sin\phi + \sin\theta) = n\lambda,$$

where the letters have the same meaning as in Art 59 If the incident wave-front coincides with the plane of the grating, $\phi = 0$ and θ becomes the deviation If the grating is used in minimum deviation (Art. 59) $\phi = \theta$ and the deviation is 2θ The deviation being capable of very accurate determination, the wave-length is found directly, when the distance between the lines of the grating is known The measurement of that distance, which is easy enough so far as the standards of length in common use may be trusted, becomes very difficult where high accuracy is required The limit of accuracy is reached when the small variations of e which always exist in different parts of the grating become appreciable, because the above equation does not hold when we substitute for e its average value and for θ the deviation of the brightest portion of the diffracted image formed by a grating in which e is not the same throughout. The careful measurements of wave-lengths by Ångström were sufficient until the increased resolving powers of modern gratings came to be used Rowland, feeling the necessity of more accurate determination to express his own results, initiated a series of measurements by Dr Bell †, and finally combining the best results fixed on $\lambda = 5896{\cdot}156$

* "On Concave Gratings for Optical Purposes", *Phil Mag* xvi p 197 (1883) See also J S Ames' "The Concave Grating in Theory and Practice", *Phil Mag.* xxvii p 369 (1889)

† *Phil Mag* Vol xxv (1888)

as the value of the wave-length of the least refrangible yellow sodium line in air at 20° C and 760 mm. pressure The unit here is the Xth metre or one metre divided by 10^{10}. This unit has been found very convenient in spectroscopic work and is sometimes called the Ångstrom unit (A U)

Concave gratings are not suitable for direct measurements of wave-lengths, but once a standard value has been obtained for one wave-length, they are useful in fixing the value of others in relation to that standard They perform this portion of the work better than plane gratings because their images need no focussing by lenses, which always show a certain amount of chromatic aberration The spectra of different orders have coincident foci. In observing the solar spectrum if one Fraunhofer line in the second order spectrum is found to overlap exactly another line in the third order, the ratio of wave-lengths of these two lines must be as 2 : 3 Rowland* has made free use of such coincidence to determine the standards of wave-lengths in different parts of the spectrum

We give some of these standards arranged according to the nature of the source

TABLE VI.

Cadmium	6438 680 × 10^{-8} cms	Hydrogen	6563 054 × 10^{-8} cms.
	5086 001		4861·496
	4800·097	Lithium	6708·070
	4678 339		6103 812
Magnesium	5711 374	Zinc	4810 725
	5528 672		4722 339
	5183·791		4680 319
	5172 866		
	5167 488		
	4703 249		
	4571 281		

The wave-lengths are measured in air at the same temperature and pressure as that of the Sodium standard The wave-lengths of Hydrogen are those observed in the solar spectrum In the other cases, the wave-lengths are those obtained in the arc, which do not always coincide in the third or even second decimal place with those observed in the sun, partly because the wave-length of light emitted by an incandescent vapour varies slightly with its pressure

* *Journal of Astronomy and Astrophysics*, Vol XII p 321, *Phil Mag* XXVII p. 479 (1889)

and partly because few lines are sufficiently sharp at atmospheric pressure to allow the third place of decimals to be accurately measured

Michelson was able, with the help of his interference method, to compare directly the wave-lengths of the Cadmium lines with the French standard of length The Cadmium lines were chosen because they were found to be very homogeneous when the metal was volatilized in vacuo It is impossible to give here a detailed account of the method used by Michelson, and a rough description possesses only little value The original papers should be consulted by anyone interested in this very beautifully devised and executed piece of work* The results for the Cadmium lines are for a temperature of 15° C and a pressure of 760 mm

Cadmium (Michelson).

$$\lambda = 6438\ 4722 \times 10^{-8} \text{ cms}$$
$$5085\ 8240 \quad ,,$$
$$4799\ 9107 \quad ,,$$

These results are less by about 19 units than those of Rowland which are based on Dr Bell's measurements The discrepancy has not yet been cleared up and can only to a small extent be explained by a difference in pressure Taking Michelson's red Mercury line as standard, Messrs Fabry and Perot have determined the lines of some other metals in vacuo Owing to the facility with which a Mercury arc in vacuo may be used to give homogeneous light, the wave-lengths of the principal lines of Mercury as given by these authors may be quoted

Mercury (Fabry and Perot)

$$\lambda = 5790\ 6593$$
$$5769\ 5984$$
$$5460\ 7424$$
$$4358\ 343$$

These being based on Michelson's standard are liable to the above systematic differences from Rowland's scale

If the wave-lengths in vacuo are required, all the above numbers must be divided by the refractive index of air, which is about 1 00028 in the visible part of the spectrum The resulting corrections amount to $1\ 90 \times 10^{-8}$ for a wave-length 7×10^{-5}, 1 37 for $\lambda = 5 \times 10^{-5}$ and 85 for $\lambda = 3 \times 10^{-5}$

68. Historical. Joseph Fraunhofer (born March 6, 1787, died June 7, 1826) was engaged from an early age in a glass manufacturing works, and became specially interested in the construction of telescope

* Details are given in the *Mém. du Bureau international de poids et mesures*, xi (1895), *C R.* cxvi p 790 (1894)

lenses He recognized the fact that their improvement, especially regards achromatism, depended on an exact determination of refracti indices, and that the chief difficulty in that determination lay in t difficulty of obtaining homogeneous radiations which could serve standards The sodium flame was made to serve as one kind radiation, and in using sunlight he discovered that nature had plac standard radiations at his disposal The spectrum of the sun was se to be traversed by dark lines—now called Fraunhofer lines—whi marked the position of homogeneous radiations by a deficiency radiance just as well as could have been done by an increase in it T the earlier observation by Wollaston, not much importance had be attached, because it had not then been recognized that their positi was invariable and independent of the mode of observation A fe years before his early death, Fraunhofer was led to the study diffraction effects and constructed the first gratings, by stretchi fine wires between two screws having narrow threads, and also l ruling lines with a diamond point on a glass surface He used the gratings for the determination of the wave-length of the princip Fraunhofer lines

Table VII gives in A U of 10^{-8} cms the wave-lengths as obtaine by Fraunhofer and subsequent observers

TABLE VII

Solar line	Fraunhofer 1823	Ångstrom 1868	Rowland 1887
C	6561	6562	6563
D	5890	5892	5893
E	5268	5269	5270
F	4859	4861	4861
G	4302	4307	4308
H	3963	3968	3969

Not much progress could be made in improving the accuracy wave-length determination until the manufacture of gratings wa improved Those made by Nobert towards the middle of last centur obtained considerable reputation, and Ångstrom (born Aug 13, 181 died June 21, 1874, in Upsala) constructed an Atlas of the Sola Spectrum with one of Nobert's gratings, which for a considerable tim remained the standard to which all wave-lengths were referred

Lewis Morris Rutherford, an amateur astronomer, and lawyer b profession, ruled gratings, by means of an automatically actin

dividing engine, which were considerably better than any previous ones He was the first to rule gratings on metal, which being softer than glass did not destroy the ruling edge of the diamond to the same extent Most of his gratings were made about the year 1880

Rowland (born 1848, died April 16, 1901) effected still greater improvements An essential portion of a machine intended to rule gratings is the screw, which should be as free from errors as possible Slight accidental displacements of the lines, so long as they are not systematic, and especially not recurring periodically, are not of serious importance Rowland's first achievement consisted in the making of a screw more perfect than any made before The following passage taken from his article on "Screw" in the *Encyclopaedia Britannica* gives an idea of the method he adopted

"To produce a screw of a foot or even a yard long with errors not exceeding $\frac{1}{1000}$th of an inch is not difficult Prof Wm A Rogers, of Harvard Observatory, has invented a process in which the tool of the lathe while cutting the screw is moved so as to counteract the errors of the lathe screw The screw is then partly ground to get rid of local errors But, where the highest accuracy is needed, we must resort in the case of screws, as in all other cases, to grinding A long solid nut, tightly fitting the screw in one position, cannot be moved freely to another position unless the screw is very accurate If grinding material is applied and the nut is constantly tightened, it will grind out all errors of run, drunkenness, crookedness, and irregularity of size The condition is that the nut must be long, rigid and capable of being tightened as the grinding proceeds, also the screw must be ground longer than it will finally be needed so that the imperfect ends may be removed "

Rowland's discovery of concave gratings, their perfection, and some of the work accomplished by them, have already been described. His measurements and maps now form the standard to which all wave-lengths are referred

CHAPTER VII.

THE THEORY OF OPTICAL INSTRUMENTS

69 Preliminary discussion. There is a limit to the pow of every instrument, due to the finite size of the wave-length of ligl According to the laws of geometrical optics, the image of a st formed in a parabolic mirror should be a mathematical point, a if this were the case the sole consideration to be attended to in t construction of optical instruments would be the avoidance aberrations. According to the wave theory of light, however, t image of a point source is never a point, however perfect t instrument may be in other respects, and the longer the wave-leng the more does the light spread out sideways from the geometric image It is therefore useless to try to avoid aberrations beyond certain point, and it becomes a matter of primary importance define the natural limit of the power of an instrument, so as to able to form a clear idea as to how far the optician may usefu spend labour in the refinement of his surfaces

Let a wave divergent from a point source A (Fig 89) be limited an aperture SS' in a scree and let the light transmitt through this aperture be further obstructed in passage by any perforat screens, but pass entire through lenses, or be flected or refracted in a manner until ultimately t

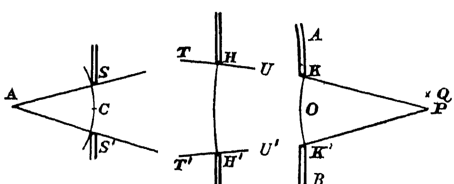

Fig 89

wave surfaces become portions of spheres concave towards a point It will be necessary to calculate the amplitude in the light in t neighbourhood of P, and a preliminary proposition will help to simpl the problem. Trace the rays AS, AS', limiting the beam, accordi to the laws of geometrical optics, and let TU, $T'U'$ be portions these rays. Place a screen at KK' with an aperture just sufficie

to enclose these rays all round, or in other words, let the edge of the aperture KK' coincide with the geometrical shadow of the opaque portions of the screen SS' The proposition to be proved is, that the introduction of this screen does not alter the distribution of light in the neighbourhood of P, and that the screen SS' may now be removed, leaving all the amplitudes near P as they were The truth of the proposition depends on the fact that all portions of the wave surface passing through KK' contribute equally to the amplitude of P, as P being a point of convergence of the rays, its optical distance to any point of KK' is the same The screen KK' obliterates only the waves which have spread out laterally before they have reached the plane of the screen. The portions so obliterated form a very small fraction of the light forming the image at P which is due to the combined action of the complete wave. The same is true for the resultant amplitude at Q so long as the aperture KK' only contains a small number of Fresnel zones drawn from Q as centre

In order that students should not be misled to apply this proposition erroneously, we may take an example where it does not hold.

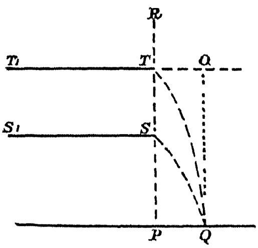

Fig. 90

SS_1 (Fig 90) is a screen limiting a parallel beam of light, RS being the edge of the geometrical shadow SS_1 cannot in this case be replaced by a screen TT_1 placed so as to touch the same limiting rays, because tracing Fresnel zones from Q backwards, the locus of the division between two zones is a parabola (Art 51) Such a parabola QS will trace the limiting zone for the screen SS_1, while if this were replaced by TT_1 the limiting curve would be a different parabola QT. If the angular space TQS includes an odd number of zones, the change of position of the screen from SS_1 to TT_1 would cause a difference in amplitude equal to that of a complete zone, so that a maximum of light might be changed into a minimum or vice versa

70. Image formed by a Lens. It is convenient to imagine the beam to be now limited by a diaphragm just inside the lens which concentrates the light at F The wave-fronts are then circles with F as centre If $D=2R$ is the diameter of the lens, ρ the distance of any point P from F, and $f=OF$,

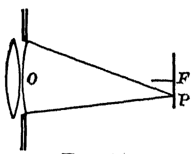

Fig. 91.

$$AP^2=f^2+(R+\rho)^2,$$
$$BP^2=f^2+(R-\rho)^2;$$
$$\therefore\ AP^2-BP^2=4R\rho,$$

and ρ being very small compared with f,

$$AP - BP = \frac{2R\rho}{f} \qquad \cdot(1)$$

If we were only to consider rays in the plane of the paper, then light at P would be destroyed by interference if

$$\frac{2R\rho}{f} = n\lambda \qquad (2),$$

and bands of maximum brightness would appear where $\frac{2R\rho}{f} = (n + \frac{1}{2})\lambda$, and if we imagine the figure to revolve round the axis OF of the lens, the luminous appearance in the plane through F, at right angles to the axis, would be a luminous disc fading outwards until the intensity becomes zero when $\rho = f\lambda/D$ This disc would be surrounded by dark and bright rings, the brightest parts of the rings corresponding to the distance $\rho = (n + \frac{1}{2}) f\lambda/D$

Owing to the rays which do not lie in the plane of the paper the destruction of light takes place at a distance somewhat greater from F than that given by the above approximate calculation.

Sir George Airy* was the first to solve the problem of the distribution of light in the diffraction image of a point source His solution depends on the summation of a series Lommel gave the solution in terms of Bessel functions The main effect is obtained more simply by the above elementary considerations The diffraction image is a disc surrounded by bright rings, which are separated by circles at which the intensity vanishes

If we write

$$\rho = m\frac{f\lambda}{D} \qquad (3),$$

the values of m for the circles of zero intensity are given in the following Table They differ very nearly by one unit, but instead of being integers, as the approximate theory would indicate, approach a number which exceeds the nearest integer by about one quarter

TABLE VIII *Dark rings*

Order of ring	m	Total light outside dark circle
1	1 220	161
2	2 233	·090
3	3 238	·062
4	4 241	048
5	5 243	039
6	6 244	032

* *Trans. Camb Phil Soc*, v p. 283 (1834)

The third column in the above table gives the amount of light lying outside each ring The first number 161 indicates that 839 of the total light goes to form the central disc while the difference between the first and second number gives the fraction of the total light which forms the first ring These differences are put down in Table IX which is mainly intended to give the values of m for the circles of maximum illumination and the corresponding intensities The third column contains the intensity at the maximum in terms of the central intensity, while the fourth column gives the fraction of the total light which goes to form the central image and each successive ring

TABLE IX *Bright Rings*

Order of disc or ring	m	Maximum intensity in terms of central intensity	Fraction of total light in disc or ring
1	0	1	839
2	1 638	01745	071
3	2 692	00415	028
4	3 716	00165	015
5	4 724	00078	009
6	5 724	00043	006

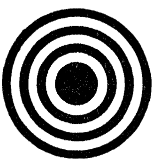

Fig 92

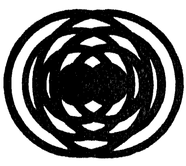

Fig 93

Fig 92 gives in diagrammatic form the relative sizes of the central disc and the first three rings Fig 93 shows the images of two sources of light placed at such a distance apart that the centre of the bright disc of one falls on the first dark ring of the other.

71. Resolving Power of Telescopes. It has long been known to all astronomers working with high powers, that the image of a star in a telescope has the appearance roughly represented in Fig 92, and it is a matter of experience that a close double star may be recognized as such when the relative position of the stars is not closer than that represented by Fig 93 This allows us to calculate the angular distance between the closest double star which the telescope can recognize as such

The radius of the first dark ring being ρ and the focal length of the telescope being f, the angle θ subtended at the centre of the object

glass by two stars which occupy such a position that the centre of the diffraction image of one falls on the first dark ring of the other is ρ/f, which by (3) gives

$$\theta = 1\ 22\ \lambda/D \quad . \quad . \quad . \qquad (4).$$

This is equal to the angular distance between the stars when they are on the point of resolution. No subsequent refraction of light through lenses can increase this angle The images may be enlarged but the rings and discs are always enlarged in the same ratio This is an important fact which may be more formally proved in this way I the rays crossing at any point o the diffraction image Q (Fig 94 are brought by a lens or system o lenses to cross again at a point Q the optical distance from Q to Q along all paths must be the same, and hence the retardation of phas between any two rays at Q is accurately reproduced again at Q'. I there is neutralization at Q, there must also be neutralization at Q As Q' is the geometrical image of Q, the diffraction pattern in th plane of Q' must be the geometrical image of the diffraction pattern i the plane of Q Our result may therefore be applied to eye observa tions through a telescope, the plane of Q' representing the plane o the retina

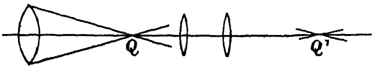

Fig 94

It appears from the above that the power of a telescope to resolv double stars is proportional to the diameter of the lens This is a resu of the wave-theory of light, for if the rays were propagated by the law of geometrical optics, the size of the object glass would not enter in the question, while the angular separation due to greater focal lengt could be increased at will by using a magnifying arrangement. W also see that the smaller the wave-length, the more nearly are the law of geometrical optics correct

To resolve stars at an angular distance of 1 second of a ($4\ 84 \times 10^{-6}$ in angular measure), we should for $\lambda = 5 \times 10^{-5}$ requi a linear aperture of

$$D = 1{\cdot}22 \times \frac{5 \times 10^{-5}}{4\ 84 \times 10^{-6}} = 12\ 6 \text{ cms}$$

Hence the angular distance θ in seconds of arc, which an obje glass of diameter D can resolve, is

$$\theta = \frac{12\ 6}{D}$$

The Yerkes telescope with an aperture of 100 cms. should be al therefore to resolve two stars at a distance of one-eighth of a second arc This calculation is based on the supposition that the whole of t light which passes through the telescope enters the eye. By a wc

known law which will be proved in Art 75, the magnifying power of a telescope is equal to the ratio of the widths of the incident and emergent beams If the width of the emergent beam is greater than the greatest width p which is capable of entering the pupil of the eye, the full aperture is not made use of Hence to obtain full resolving power the magnifying power of a telescope should be not less than D/p If it is less, the rays entering the outer portion of the telescope lens do not enter the eye at all and may as well be blocked out altogether, thus reducing the aperture to its useful portion

72. Resolving Power of the Eye We may apply equation (4) to the case of two stars or other point sources being looked at directly by the eye An apparent complication arises owing to the fact that the wave-length of light in the vitreous humour, which is the last medium through which it passes, is not the same as the wave-length in air, but this makes no difference provided that we take for p the width of the beam as above defined Let Fig 95 represent diagrammatically a beam of light entering the media of the eye If a plane wave-front passes through an aperture AB of such size that the beam passing through it may just enter the pupil of the eye, the first dark ring of the diffraction images passes through Q when the difference in optical lengths from A to Q exceeds by 1 22λ that from B to Q Also a wave-front parallel to $W'F'$ has the centre of its diffraction image at Q when the optical distance from all points of its plane to Q is the same, hence AT must be equal to 1·22λ, and the angle between AB and $W'F'$ is measured by AT/AB or

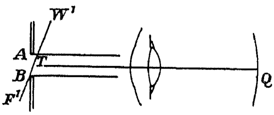

Fig 95

$$1\ 22\frac{\lambda}{p}$$

Here λ is the wave-length measured in *air*

The width of pupil is variable, but with light of medium intensity such that p is about 3 mm (the actual opening of the pupil will be less, owing to the convergence produced by the cornea), two small point sources of light should be resolvable by the eye when at an angular distance of 42″ Helmholtz gives for the experimental value of the smallest angular distance perceptible by the eye the range between 1′ and 2′, which would show that with full aperture of the pupil, our sense of vision is limited rather by the optical defects of the eye and physiological causes than by diffraction effects.

73. Rectangular Apertures. If the surface of the telescope is covered by a diaphragm having a rectangular aperture, the distribution of light is more easily calculated, and may be expressed accurately in a

simple form Take the axis of x and y in the focal plane of the lens and parallel to the sides of the aperture and let the length and width of the aperture in the direction of x and y be a and b respectively

When b is large we obtain the case investigated in Art 53, where the intensity was found to be proportional to $(\sin^2\alpha)/\alpha^2$ where $\alpha = \pi ax/f\lambda$. If a is large, the expression must be proportional to $(\sin^2\beta)/\beta^2$ where $\beta = \pi by/f\lambda$. We can satisfy both conditions if in the general case we take the intensity to be proportional to $\sin^2\alpha\sin^2\beta/\alpha^2\beta^2$ The constant to be applied may be found by considering that if a and b are very small, the amplitude must by Art 46 be equal to $ab/f\lambda$ if the incident beam has unit intensity Hence for the complete expression we obtain

$$I = \frac{a^2b^2}{f^2\lambda^2}\frac{\sin^2\alpha}{\alpha^2}\frac{\sin^2\beta}{\beta^2} \qquad (5)$$

The most important case we shall have to consider is that where the source is a luminous line parallel to the direction of y A point of the luminous line at a distance y' from its central point causes an intensity at the central point of the image which may be obtained from (5) by making the value of y contained in β equal to $y'e/e'$ where $2e'$ is the length of the source and $2e$ that of its image Hence the total intensity at the point $y = 0$ is proportional to

$$\int_{-e'}^{+e'} I dy' = \int_{-e}^{+e} I dy$$

If e is large, we may substitute infinity for the limits, and as

$$\int_{-\infty}^{+\infty} \frac{\sin^2 v^2}{v^2} dv = \pi$$

it follows that the intensity in the image is proportional to

$$I' = \int_{-\infty}^{+\infty} I dy = \frac{a^2b}{f\lambda}\frac{\sin^2\alpha}{\alpha^2},$$

where α has the same value as before It follows that the total amount of energy which is transmitted in unit time through a small surface s of the image is $\kappa s I'$, where κ is a constant which may be determined as follows If a ribbon of unit width be cut out transversely to the image, the total amount of energy transmitted through the ribbon is

$$\kappa\int_{-\infty}^{+\infty} I'dx = \kappa ab$$

If E denote the amount of light from unit length of the source transmitted through unit surface of the first lens, and m the magnifying power, the total amount of light per unit length of the image is $mEab$ Hence $\kappa = mE$

74. Luminous Surfaces. The image of a surface bounded by a straight edge may be calculated from the above Dividing the

surface into narrow strips parallel to one of the edges, each strip will have a diffraction image according to Fig 70, and at each point of the

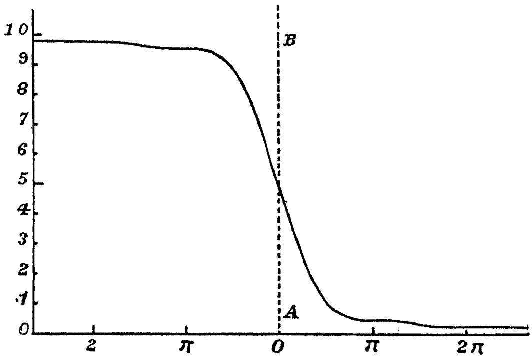

Fig 96

image we should have to add up the effect due to each strip. It is easy to see that at the geometrical image of the edge, the intensity is half that observed at some distance inside the edge, where the illumination is uniform, for when two similar surfaces are placed against each other with their edges AB in contact, a uniformly illuminated sheet is obtained, and each half must contribute equally to the illumination at the dividing line The intensity at other points can only be expressed in the form of definite integrals or calculated by means of a series The intensities are plotted in Fig. 96. The dotted line AB marks the edge of the geometrical image of the surface The intensity at that point is 5, and falls off rapidly towards the outside of the image

When a telescope is used to examine such a surface as the moon, it is not a question of separating two luminous points or sharply defined surfaces, but rather of interpreting changes of luminosity in a continuously varying surface Details which are as near together as two stars when at the point of optical separation may be indistinguishable on an illuminated surface If we double that distance the central diffraction bands stand altogether clear of each other, and hence the angular distance between two points should be equal to $2{\cdot}44\lambda/D$, if there is to be no overlapping at all. The edge of the image of a luminous surface is not bounded by alternately bright and dark fringes, and there is no definite boundary at which the image of the surface can be said to end

For a given distance from the geometrical edge the intensity is less than at the same distance from the image of a narrow aperture. Hence, as has been pointed out by Wadsworth, the images of two surfaces

y be put closer together than the images of the slit without their iges becoming confused

The points marked π and 2π on the horizontal line of Fig 96 resent the places where the first two minima of light would occur in image of a narrow slit coincident with AB

75. Illumination of the image of a luminous surface The resultant energy which leaves a luminous surface is the same in all directions for equal cross-sections of the beam. As with a given small surface S Fig 97 the cross-section of the beam varies as the cosine of

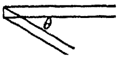

Fig 97

direction angle, θ the intensity of radiation sent out by a surface S proportional to $\cos\theta$, but for small inclinations to the normal, we y take the radiation to be independent of the direction If an ige of a surface S is to be formed, the illumination of the image st be proportional to the amount of light which the luminous face sends through the optical system If all the light which sses through the first lens passes also through the other lenses, s is proportional to the surface S, and to the solid angle ω subtended the lens at a point of S We may therefore write for the light ssing through the optical system $IS\omega$, where I solely depends on the ninosity of the surface. If s is the size of the image of S, and if image is such that the illumination is uniform, the brightness of image is equal to $IS\omega/s$

We shall first consider the case that the linear dimensions of s are ch that the diameter of the diffraction disc may be neglected in mparison with it, so that we may find the relation between S and s the laws of geometrical optics.

Let LL' and MM' be the wave-fronts diverging from P and converging to the image Q respectively, and imagine a second wave-front RR' slightly inclined to the first, to diverge from P' If $PO=P'O$, the second wave-front may be obtained by turning LL'

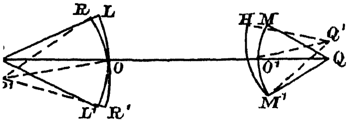

Fig 98

out O through a small angle θ The optical length from P' to L has en increased by the change, by the quantity RL, and the optical ngth from P' to L' has diminished by the same amount The tical lengths from L to M and $L'M'$ have not been altered (Art. 23) ence if Q' is the image of P' the optical length $M'Q'$ must differ from Q' by $2RL$, the total length from P' to Q' being the same whether easured by way of LM or by way of $L'M'$ It follows that to obtain ' we must turn round the wave-front $M'M$ through such an angle at $HM=2RL$ If D is the width of the beam at LL' and d the

width at MM', the angles POP' and QOQ' are $2RL/D$ and MH/d respectively, and are therefore in the inverse ratio of $D \cdot d$ It follows that

$$\frac{PP'}{QQ'} = \frac{d \times PO}{D \times O'Q} \qquad (6)$$

If a square of surface S and sides P, P' is formed in a plane at right angles to OO', its image S will be a square with Q, Q' as sides, hence

$$\frac{S}{s} = \frac{d^2}{D^2} \frac{PO^2}{O'Q^2}.$$

The solid angle (ω) of the beam entering the first lens is $\pi D^2/4PO^2$, and the solid angle (ω') of the beam converging to Q is $\pi d^2/4O'Q^2$.

Hence the illumination per unit surface of s is

$$\frac{IS\omega}{s} = I \cdot \frac{d^2}{D^2} \frac{PO^2}{O'Q^2} \cdot \frac{\pi D^2}{4PO^2}$$

$$= I \frac{\pi d^2}{4O'Q^2} = I\omega' \qquad (7)$$

Before discussing the last equation, we note two interesting results which have incidentally been obtained in the investigation

QQ' is inverted as compared with PP' and this must always be the case according to the construction when the limiting ray MQ is the continuation of the ray PL on the same side of the axis, but if the rays have crossed once or an odd number of times between O and O', so that the ray PL becomes the ray $M'Q$, we should have to turn round the ultimate wave-front MM' in the opposite direction in order to equalize the optic lengths of the extreme rays, and the image would then be erect.

The ratio of the angles $QO'Q'$ and POP' becomes the magnifying power (m) of the arrangement, when, as in a telescopic system, the incident and emergent beams are both parallel, hence

$$m = \frac{QQ'}{QO'} - \frac{PP'}{PO}$$

and by (6),
$$= \frac{D}{d},$$

which proves the proposition which has already been made use of in Art 71

The theorem, defined by equation (7), that the brightness of a luminous surface is determined by the solid angle of the converging pencil which forms the image, is of fundamental importance We may derive three separate conclusions from it (1) The apparent brightness of a luminous surface looked at with the naked eye is independent of its distance from the observer. (2) No optical device can increase the

apparent brightness of a luminous surface above what it is when the surface is looked at with the naked eye (3) When looked at through a telescope the brightness of a surface is independent of magnifying power up to a certain limit, and above that limit, the brightness varies inversely as the square of the magnifying power

The first of these propositions depends on the fact that when looked at with the naked eye, the solid angle on which the brightness depends, is determined solely by the width of the pupil, and the dimensions of the eye, and, independently of casual changes of the pupil, is constant. Hence the brightness of the solar disc is the same when looked at from the furthest or from the nearest planet The total amount of luminous radiation no doubt diminishes as the distance increases, but the apparent size of the disc diminishes in the same ratio, and hence follows the equality of the amount of light *per unit surface* of the image on the retina Elementary considerations are sufficient to show that the apparent size of the image of a surface varies inversely as the square of the distance and that illumination is therefore constant, but the second and third of the above propositions are not quite so obvious Imagine a surface, *e g* the moon, looked at through a telescope having an aperture of diameter D So long as the magnifying power is less than D/p, where p is the diameter of the pupil, the width of the beam entering the eye is p, and the solid angle ω' is the same as if the moon were looked at with the naked eye The moon would therefore appear to be of exactly the same brightness in the two cases, if there were no loss of light by reflexion and absorption in the optical media of the telescope, but in no case can the moon appear brighter through the instrument When the magnifying power (m) is greater than $m' = D/p$, the width of the emergent pencil is $d = D/m$ and the solid angle ω' is reduced in the ratio d^2/p^2 or D^2/p^2m^2 Hence for magnifying powers greater than m', the brightness is reduced into the ratio m'^2/m^2 In observing luminous surfaces, therefore, through a telescope, we may apply magnifying powers up to D/p without loss of brightness through magnifying power, but we do not make use of the full aperture in that case at all, so that to obtain the full resolving power and full illumination, the magnifying power should be D/p Taking the aperture of the pupil to be 3 mm. this would give a magnifying power of $3\frac{1}{3}$ for each centimetre or about nine per inch of aperture There is, however, an advantage in using somewhat higher magnifying powers, as the outer portions of the crystalline lens do not assist the definition on account of aberration Most eyes see objects therefore more distinctly when the size of the pupil is reduced to about 2·5 mm which would give a magnifying power of 4 for each cm of aperture. With greater magnifying powers, there is no gain in definition and there is loss in brightness It should be noted that in all cases so far

considered, the brightness of the image does not in any way depend o the focal length of the lens. It is otherwise when telescopes are use for photographic purposes The solid angle ω' on which the brightne depends, varies in this case with $(D/f)^2$, D being the diameter an f the focal length A short focus lens of large diameter is therefo of considerable advantage in these cases

76. Brightness of Stars. The above results apply only so lo as the size of the image of a surface is large compared with the size of t diffraction image Other considerations regulate the brightness of t image of a star The diameter of the diffraction image of a star h been shown to be inversely proportional to the aperture Wh looked at with the naked eye, or through a telescope of low ma nifying powers, the diameter of the disc is determined by the wid of the pupil, and the brightness varies in that case as the amou of light which enters the eye If the magnifying power is D/p, t amount of light collected by the lens is D^2/p^2 times that collected the unaided eye. Hence the illumination of the image of a star var as the square of the effective aperture of the lens, so long as t magnifying power is adjusted so as to be equal to D/p If less th that, we must imagine the unused portions of the lens to be cover and the aperture reduced to its "effective" portion When the ma nifying power is D/d, d being smaller than p, the linear size of t diffraction image is increased in the ratio p/d, so that the brightne now will vary as D^2d^2/p^4 For star images as well as for finite surfa there is therefore loss of light without gain in definition, when t magnifying power is increased above a certain value Astronome frequently, however, use a higher power than that which according the above should give the best results The reason is physiologic Increased size of the diffraction images, even though without incre of *optical* definition, and accompanied by loss of light, assists facility observation, and increases therefore what may be called *physiologi* definition

The increase in the number of stars visible through telescopes easily accounted for While the general brightness of the sky rema the same, or taking account of loss of light by reflexion and absorpt is diminished, when an instrument is used, the brightness of a sta increased fifty times by the use of an opera-glass having an apert of not quite an inch, and the largest telescopes allow the light wh enters the eye from each star to be increased 100,000 times. It is surprising then that the number of stars which become visible telescopes is considerably increased

The planets occupy an intermediate position between the moon a the fixed stars. When looked at with the naked eye, the diamete

the image on the retina is less than that of the diffraction disc, but with Venus, Jupiter and Saturn it is only a few times smaller The use of a telescope having an aperture up to ten times the diameter of the pupil would when applied to planets be accompanied by an increase of brightness, but after that point is reached, the larger planets named would behave like bodies of finite surface An increased aperture would only act by allowing higher magnifying powers to be used

77 Powers of Spectroscopes. A spectroscope may be used for two different purposes In the majority of cases, it serves to examine the radiations of a luminous source, by separating the radiations if homogeneous, or giving us the distribution of intensity, if non-homogeneous But another not less important function of the spectroscope is to produce homogeneous light. By allowing the spectrum formed by a source of white light to fall on a screen with a narrow slit placed so that only rays very near those of a certain wave-length pass through the slit, we obtain a source of nearly homogeneous radiations The power of a spectroscope may, irrespective of the source of light used in conjunction with it, be defined as its power to produce homogeneous light Its power to separate two homogeneous radiations of nearly the same wave-length, which may be called its resolving power, is found to depend on exactly the same conditions as its power to produce homogeneous radiations The problem as regards gratings has already been to some extent dealt with in Art 62. The radiations sent out by luminous vapours are often very complicated and sometimes consist of one or more nearly homogeneous radiations lying close

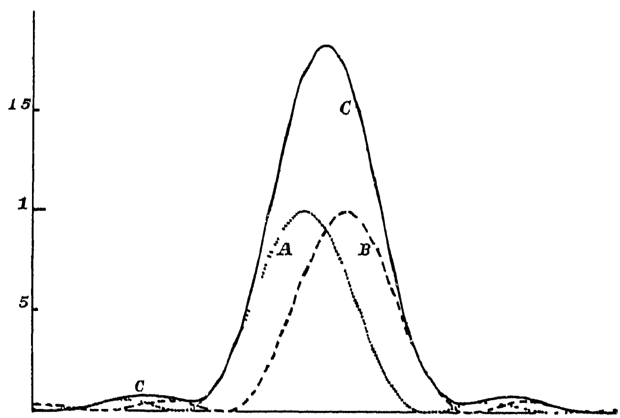

Fig 99

together Consider a source of light sending out waves, the lengths of which, λ_1 and λ_2, differ but little from each other If the light, after

passing through a slit, rendered parallel by a "collimator," falls on a grating, and is then collected by a lens, two images will be formed at the focus The diffraction image of each is of the same kind as that of a luminous line in a telescope, the object glass of which has been covered by a screen with a rectangular aperture, because the grating itself causes the cross section of the effective beam to be rectangular If the difference between λ_1 and λ_2 is very small, there is a considerable overlapping, and what the eye perceives is the sum of the intensities at each plane of the two images In Fig 99 the curves A and B show the distribution of intensity of the two separate slit images, while C gives the sum of the intensities The combined curve is so nearly equal to the curve of the image of a single slit that the eye could not realize that the light is made up of two different wave-lengths The two lines are not in that case "resolved" Fig 100 gives the combined intensities of the same two lines, when placed three times as far apart, and at such a distance that the maximum

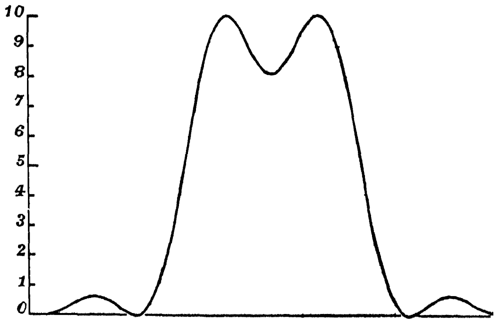

Fig 100

intensity of one image falls on the first minimum of the other. The curve shows in this case a decided dip in the middle between two maxima The intensity at the lowest point is very nearly 8 of the intensity at the maximum, and the eye clearly perceives that it is not dealing with a homogeneous radiation. The natural interpretation of a distribution of intensity such as that indicated in Fig 100 is that the radiation consists of two homogeneous radiations having wave-lengths corresponding to the positions of the maxima The two lines are then said to be "resolved," but it is of course possible, and frequently the case, that the radiation is of a more complicated character Not until the distance between the two lines is about double that indicated in the figure, do they stand altogether clear of each other. According to Art. 59, two wave-lengths λ and λ' have the relative position indicated by Fig 100, if

$$nN\frac{\lambda' - \lambda}{\lambda} = \pm 1,$$

N being the number of lines on the grating and n the order of the spectrum In order just to resolve this difference in wave-lengths $\delta\lambda$ must be such that

$$\frac{\delta\lambda}{\lambda} = \pm \frac{1}{Nn}$$

The smaller $\delta\lambda$ the more powerful is the instrument for the purpose of separating double lines, and we call as already pointed out Nn the "resolving power" of the spectroscope There is something arbitrary in this definition, as the dip in intensity necessary to indicate resolution is a physiological phenomenon, and there are other forms of spectroscopic investigation besides that of eye observation In a photograph or a bolometer, the test of resolution is different It would therefore have been better not to have called a double line "resolved" until the two images stand so far apart, that no portion of the central band of one overlaps the central band of the other, as this is a condition which applies equally to all methods of observation This would diminish to one half the at present recognized definition of resolving power. Confusion would result from a change in a universally accepted definition, but it should be understood that if R is the resolving power, a grating spectroscope will completely separate two wave-lengths differing by $\delta\lambda$ only when

$$\frac{\delta\lambda}{\lambda} = \frac{2}{R}$$

78. Resolving Power of Prisms It has been proved in Art 24 that if in a parallel beam of light, two sets of waves are originally superposed, the angle θ between the two beams after passing through a prism is

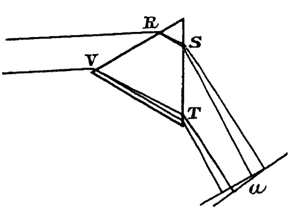

Fig 101

$$\theta = (\mu_2 - \mu_1) \frac{VT - RS}{a},$$

where μ_2 and μ_1 are the two refractive indices, and a the width of the beam after emergence The difference $VT-RS$, for which we write t, is the difference between the paths inside the prism of the extreme rays of the beam. If the prism is placed so that one of the extreme rays just passes by the edge, $RS=0$ and t will measure the greatest thickness of the dispersive material through which the ray has passed It is easy to extend the investigation so as to include any number of prisms If $T=\Sigma t$ measures the difference in aggregate thickness of the material through which the extreme rays have passed,

$$\theta = (\mu_2 - \mu_1)\, T/a \qquad (8),$$

the material here being considered the same for all prisms

This expression leads at once to the resolving power of prism spectroscopes The beam passing through the prisms having a rectangular cross-section, the angle subtended at the centre of the focussing lens by the half width of the central diffraction band is λ/a (Art. 53), hence with the definition of resolution of the last article, we have for the two wave-lengths at the point of separation, $\theta=\lambda/a$ or

$$\frac{(\mu_2-\mu_1)}{\lambda}=\frac{1}{T}$$

The ratio $(\mu_2-\mu_1)/(\lambda_2-\lambda_1)$ is the rate of increase of refractive index per rate of increase of wave-length, and may for small differences of wave-length be written $d\mu/d\lambda$. Hence for the resolving power

$$\frac{\delta\lambda}{\lambda}=1\Big/\left(T\frac{d\mu}{d\lambda}\right),$$

and

$$R=-T\frac{d\mu}{d\lambda}$$

The minus sign is necessary, as R is essentially positive and $d\mu/d\lambda$ negative

This fundamental relation, due to Rayleigh, shows that the resolving power of prism spectroscopes is proportional to the greatest thickness of the dispersive material traversed by the rays (the edges of the prisms being arranged along the path of one of the extreme rays of the beam)

The distinction between the dispersion and the resolving power is a very important one Confining ourselves to one prism, the dispersion $\theta/(\mu_2-\mu_1)$ is obtained from (8) and varies inversely with the cross-section of the beam If a prism be placed in one of the two positions A and B (Fig 102), the position A gives a greater dispersion than B, in the ratio of t_1/a_1 t_2/a_2, but the resolving powers only vary in the ratio $t_1 : t_2$. The greater dispersion is therefore not accompanied by a correspondingly greater resolving power, the reason being that the narrow beam of A, though giving greater dispersion, gives also a broader diffraction image The increased dispersion means therefore chiefly increased magnification without increased definition With ordinary flint glass $d\mu/d\lambda$ is about 1000 in the neighbourhood of the sodium line, so that one centimetre of glass is sufficient to separate the two sodium lines, the difference between their wave-lengths being very nearly equal to the thousandth part of the wave-length of one of them. When the prism is in the position of minimum deviation, the

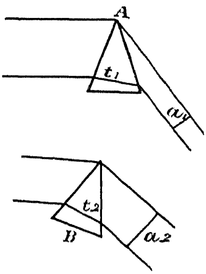

Fig. 102

length t is equal to the thickness of the base of the prism The superiority of gratings over prisms as regards resolving power is shown by the fact that the gratings in common use have about 5600 lines to the centimetre, and if ruled over a distance of 5 cms the total number of lines would be 28,000 To produce with prisms a spectroscope of resolving power equal to that of the first order spectrum of the grating, would require a thickness of 28 cms. of glass, or say 7 prisms, having a base of 4 cms each

It will be noted that the resolving power of prisms depends on the total thickness of glass, and not on the number of prisms, one large prism being as good as several small ones Thus all the prisms drawn in Fig 103 would have the same resolving power, though they would show very considerable differences in dispersion

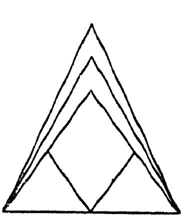

Fig 103

79. Resolving Power of Compound Prisms. The only kind of compound prism we need consider here, is the one giving direct vision Two prisms of crown glass A, A' (Fig. 104), may be cemented to a prism B of flint adjusted so that a ray of definite wave-length falling on the compound prism in the direction of its axis, passes out in the same direction It follows by symmetry that in the centre prism B, the direction of the ray must then also be along the axis The extreme thicknesses travelled through by the rays are, on one side a thickness t' in flint, and on the other side, a thickness $2t$ in crown The resolving power of such a prism is

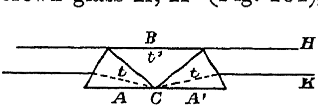

Fig 104

$$\frac{t' d\mu'}{d\lambda} - 2t \frac{d\mu}{d\lambda},$$

where $\frac{d\mu'}{d\lambda}$ refers to flint and $\frac{d\mu}{d\lambda}$ to crown. The dispersion of the crown glass is here opposed to that of the flint and the resolving powers of such compound prisms are small

80. Greatest admissible width of source. In considering the effects of interference and diffraction, we had considered the source of light to be either a point or a line, but in the actual experiment, every source has finite dimensions, and as in general it is important to secure as much illumination as possible, these dimensions are increased as much as is consistent with good definition The limit to which we can increase the dimensions of the source depends to some extent on the object we have in view. When, e g, we wish to measure accurately a spectroscopic line known to be single, we may

use a much wider slit than if we wish to see whether a line is single or double But even in the last case, there is a limit below which very little is gained by narrowing the source To determine this limit, we consider the diffraction pattern of the image of a slit Widening the slit alters the law according to which the intensity of light is distributed in the diffraction image, and it may be seen from an inspection of Fig 70 that increased width of slit means an increase in intensity which is greater for the weaker portions near the minimum than for the central portion at the maximum Computation shows however that until the geometrical image of the slit exceeds the eighth part of the width of the central diffraction band, the alteration in the distribution of light is insignificant, so that there is not much advantage in narrowing the slit beyond this limit Let light passing through the slit at O' (Fig 105) and the collimating lens $A'B'$ ultimately be focussed by the lens AB. The centre of the image being at the principal focus, the geometrical image of the slit has a width equal to the eighth part of the diffraction band when $BO - AO = \lambda/8$, O being the image of one of the edges of the slit Fermat's principle at once leads to the conclusion that if O' is the edge of the slit which has its image at O, $A'O' - B'O' = \lambda/8$ The width of the slit is then found by geometrical considerations to be $f\lambda/4D$, where D is the aperture and f the focal length of the collimator lens Writing ϕ for the angle subtended by the collimator at the slit, the greatest admissible width (d) of slit, for full definition, becomes

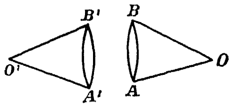

Fig 105

$$d = \frac{\lambda}{4\phi}$$

ϕ is generally about 1/16, so that the width of the slit should not be more than four wave-lengths When extreme definition is not required, we may, without seriously interfering with the accuracy of the observations, allow a difference in phase of a quarter of a wave-length between the extreme rays This would double the admissible width of slit. Two spectrum lines placed in the position of Fig 100 would show with this width of slit a diminution in intensity amounting to 10 % at the lowest point of the intensity curve, instead of 20 % which they give with an indefinitely narrow slit The resolution would be more difficult, but under favourable circumstances not impossible, as to some extent, the smaller variation in intensity is counterbalanced by increase in brightness. The above condition $A'O' - B'O' = \lambda/8$ may conveniently be expressed in a different form. Let O'' be the other edge of the slit, then by symmetry $A'O' = B'O''$ and hence

$$B'O'' - B'O' = \lambda/8.$$

We may say therefore that for perfect definition the admissible width

is determined by the condition that the distances from different points of the source and any one point of the edge limiting the beam, shall not exceed one eighth part of the wave-length In many cases this difference may be increased to one quarter of a wave-length. When put in this form the proposition is of great practical utility Thus if the bright spot at the centre of the shadow of the circular disc of diameter d is to be observed and f be the distance from the disc of a small opening, through which the light enters, we may take the linear dimensions of the opening to be $f\lambda/2D$

81. Brightness of image in the spectroscope. When we are dealing with homogeneous light, the investigation of Art 75 shows that we always get the maximum illumination when the pupil is filled with light This determines the magnifying power of the telescope, at which there is both full resolving power and full illumination. The former is lost by diminishing, while illumination is lost by increasing the magnifying power Errors are frequently committed, owing to the belief that illumination depends on the ratio of the aperture to the focal length either in the collimator or the telescope of the spectroscope This is not correct It is important, however, to consider both the height and the width of the beam as it leaves the grating or prism The prism narrows or widens the beam, unless it is in a position o minimum deviation, and with a grating the width depends on the angle at which the spectra are observed. When it is important to magnify, even at the risk of losing light, the spectroscopes have an advantage over telescopes, as by placing the prism out of minimum deviation, in such a way that the beam is narrowed, we enlarge the image in one dimension only, and it is just that lateral magnification which is required Hence the corresponding loss of light is less than it would be if the enlargement were done by a higher power eye-piece The so-called half-prisms act in this way, spectroscopes being constructed of considerable magnifying power but comparatively small resolving power by cutting the direct vision compound prisms (Art. 79 into two equal halves at right angles to the axis and using one of th halves only Light falling on the face which stands perpendicular t the axis, enters the prism with a width BC (Fig 104) and leaves with a reduced width HK.

Spectroscopic investigation has often to be applied to sources o light which are so weak that the width of the slit must be increase beyond the limits compatible with complete resolving power. To for an idea of the diminution in optical efficiency, let d be the width o the slit and ϕ the angle subtended by the collimating lens at the sli Then the difference in optical length between $A'O'$ and $B'O'$, Fig. 1((O' being one of the edges of the slit), is $\phi d/2$, and this is also tl

difference $BO - AO$ where O is the image of the edge of the slit. The centre of the diffraction band of the image of the edge of the slit being at O, the first band will extend to a point T such that $BT - AT = \lambda + \phi d/2$. The ratio of this number to λ, gives the ratio of the total width of the band to that which it would have with an indefinitely narrow slit. (All light except that in the central image is here neglected.) In order that two lines should stand absolutely clear of each other with a narrow slit, it has been found that

$$\frac{\delta\lambda}{\lambda} = \frac{2}{R},$$

and hence with a slit of width d, this ratio becomes

$$\frac{\delta\lambda}{\lambda} = \frac{2\lambda + \phi d}{\lambda R} \qquad (9).$$

The light at the centre of the diffraction image increases for narrow slits with the width of the slit, but after a certain limit is reached, further widening does not cause further appreciable increase in illumination. This may be seen with reference to Fig. 79. The light passing through a narrow opening OK is still increased by a diffraction image in a position such as that of the curve marked λ_2, but neglecting all light beyond the first minimum the limit is reached when the first two minima of the light coming from the edges of the slit coincide. This leads to the condition $A'O' - B'O' = \lambda$. For the width of the slit we then have $\phi d = 2\lambda$ and hence from (9)

$$\frac{\delta\lambda}{\lambda} = \frac{4}{R},$$

so that the optical power is reduced to one half. If we take the narrower test of resolution, for which

$$\frac{\delta\lambda}{\lambda} = \frac{1}{R},$$

the reduction in optical power is even greater.

Full resolving power is only obtained if the collimator lens is completely filled with light. Hence when the slit S is wide, and the source of light (L) is narrow, it is necessary to interpose a lens (A) as shown in Fig. 106; the angular aperture of the lens A as seen from the slit need not, in this case, be larger than the angular aperture of the collimator lens. If the diameter of the lens A is increased without change of focal length, the image of L on the slit plate becomes brighter, but the increase in light is caused by rays which do not pass through the collimator lens at all, and are therefore useless. When the slit is made so narrow that full or nearly full resolving power

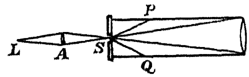

Fig. 106

is obtained, diffraction will cause the light inside the collimator to spread, so that in the arrangement as drawn in the figure a good deal of light is lost To make up for this, the diameter of A should be increased, when some of the light which with a wide slit falls on the sides of the collimator tube, will be diverted so as to pass through the collimator In the observation of star spectra, the telescope lens performs the function of the condensing lens A, and its aperture being fixed, there is necessarily a not inconsiderable loss of light due to diffraction, when full resolving power is obtained Diminishing the focal length of the collimator does not help here, because this would imply a narrowing of the slit and further spreading out, if the resolving power is to be maintained Light being of importance when star spectra are observed, it becomes necessary to build the spectroscopes so as to have a resolving power three or four times greater than that required, and to open the slit to a width of not less than $2\lambda/\phi$ and not greater than $4\lambda/\phi$

It should be noted that if a star image were perfectly steady and undisturbed by irregular atmospheric refraction, a star spectroscope should give full resolving power without any slit at all Indeed in this case, the slit could only deteriorate, but never improve the definition The tremors of the star images, due to atmospheric fluctuations, are however sufficiently serious to render a slit desirable, when high resolving powers are required

The above treatment of the subject is based on the consideration of spectra of bright lines, and cannot without modification be applied to the absorption phenomena exhibited in the spectra of sun and stars. It would lead too far to enter into this part of the subject here, but one example of the difference presented by emission and absorption spectra, may be pointed out. A perfectly homogeneous radiation could never appear as a dark line in an absorption spectrum, for the reason that an indefinitely narrow gap between two bright surfaces could not be detected by any instrument of finite resolving power.

When photographic impressions of spectra are required, the angular aperture of the lens forming an image determines the brightness of the image at the focus, and the considerations of Art. 75 may be applied.

82. Aberrations. If a wave-front approaching a point is truly spherical in shape, the amplitude at the point is as great as it can possibly be, but owing to defects in the working of the surfaces, or insufficient homogeneity of the material, perfection is never attained. From the point of view of the wave-theory of light, the so-called optical "aberrations" are dependent on the deviations of the wave-fronts from the ideal spherical shape The amount of deviation compatible with

sensibly perfect definition has been discussed by Lord Rayleigh*, who finds that if the discrepancy in phase at the focus between the extreme and central rays of the wave-front does not exceed a quarter of a wave-length, the image does not suffer appreciably Beyond that limit, there is a rapid deterioration of definition Lord Rayleigh also gives the following important applications of this result If in a telescope supposed to be horizontal, there is a difference in temperature between the stratum of air along the top and that of the rest of the tube, the wave-fronts are distorted along the top of the tube The final error of optical length of the extreme rays is $l\delta\mu$, where l is the length of the tube and $\delta\mu$ the alteration in refractive index At ordinary temperatures $\delta\mu$ is connected with δt the change in temperature, by the approximate relation

$$\delta\mu = -1\ 1\ \delta t \times 10^{-6}$$

If the error in optical length is a quarter of a wave-length,

$$\tfrac{1}{4}\lambda = 1\ 1\ l\delta t \times 10^{-6},$$

$$\therefore\ l\delta t = 12 \text{ if } \lambda = 5\ 3 \times 10^{-5}$$

Thus a change of temperature of 1° becomes appreciable when the length through which the temperature difference extends is 12 centimetres In a telescope tube 12 metres long, the average temperature of the air through which the different rays pass should not differ by more than 0 01

As a second example, also given by Lord Rayleigh, we may take the accuracy which is required in the working of optical surfaces If AC is the optical surface and if through imperfections in the working any portion of it is raised so as to occupy the position DF, the error in optical length is (Fig 107)

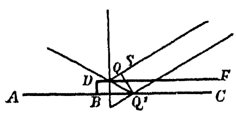

Fig 107

$$QQ' - QS = 2BD \cos\phi,$$

where ϕ is the angle of incidence and S the foot of the perpendicular from Q' to the ray reflected from Q Hence the deviation BD from the plane AC should not, over any considerable portion of the surface, exceed $\frac{1}{8}\lambda \sec\phi$, or for normal incidence, one-eighth of the wave-length.

Our result applies to the case where no change of focus is allowed in the observing telescope, but aberrations, in the sense here introduced, may often be corrected for by such change of focus as *e g* when a surface intended to be plane is slightly concave or convex

Students may, as an example, work out the greatest admissible

* *Collected Works*, Vol I. p. 428

width compatible with perfect definition of a spherical mirror, when rays parallel to the axis fall on it

83. The formation of images without reflexion and refraction. Pin-Hole Photography. An elementary experiment in Optics consists in showing the rectilinear propagation of light by projecting an image on a screen, the image being formed by rays which have passed through a narrow aperture Lord Rayleigh has shown that for small apertures, such an opening acts as well as a lens, and the discussion of the matter is here given in his own words —

"The function of a lens in forming an image is to compensate by its variable thickness the differences in phase which would otherwise exist between secondary waves arriving at the focal point from various parts of the aperture If we suppose the diameter of the lens ($2r$) to be given, and its focal length f gradually to increase, these differences of phase at the image of an infinitely distant luminous point diminish without limit When f attains a certain value, say f_1, the extreme error of phase to be compensated falls to $\frac{1}{4}\lambda$ Now, as I have shewn on a previous occasion*, an extreme error of phase amounting to $\frac{1}{4}\lambda$, or less, produces no appreciable deterioration in the definition, so that from this point onwards the lens is useless, as only improving an image already sensibly as perfect as the aperture admits of Throughout the operation of increasing the focal length, the resolving-power of the instrument, which depends only upon the aperture, remains unchanged, and we thus arrive at the rather startling conclusion that a telescope of any degree of resolving-power might be constructed without an object-glass, if only there were no limit to the admissible focal length This last proviso, however, as we shall see, takes away almost all practical importance from the proposition

"To get an idea of the magnitudes of the quantities involved, let us take the case of an aperture of $\frac{1}{5}$ inch (inch = 2 54 cms), about that of the pupil of the eye The distance f_1, which the actual focal length must exceed, is given by

$$\sqrt{\{f_1^2 + r^2\}} - f_1 = \tfrac{1}{4}\lambda,$$

so that
$$f_1 = 2r^2/\lambda$$

Thus, if $\lambda = \frac{1}{40,000}$, $r = \frac{1}{10}$, $f_1 = 800$.

"The image of the sun thrown on a screen at a distance exceeding 66 feet, through a hole $\frac{1}{5}$ inch in diameter, is therefore at least as well defined as that seen direct In practice it would be better defined, as the direct image is far from perfect If the image on the screen be regarded from a distance f_1, it will appear of its natural angular

* *Phil Mag* Nov. 1879 (Art XLII § 4).

magnitude Seen from a distance less than f_1 it will appear magnified Inasmuch as the arrangement affords a view of the sun with full definition and with an increased apparent magnitude, the name of a telescope can hardly be denied to it

"As the minimum focal length increases with the square of the aperture, a quite impracticable distance would be required to rival the resolving-power of a modern telescope Even for an aperture of four inches f_1 would be five miles"

Returning to the subject in a later paper, Lord Rayleigh discusses its application to the so-called pin-hole photography, in which the lens of a camera is simply replaced by a narrow aperture. If this aperture is too small, the image loses in definition owing to the spreading out of the waves, and on the other hand it is clear that no image can be formed, when the aperture is large There must therefore be one particular size of the opening which gives the best result The original paper* should be consulted, in which the question is treated both theoretically and experimentally The best result in general is found, when the aperture as seen from the image includes about nine-tenths of the first Fresnel zone, so that if a is the distance of the object, b that of the image from the screen and r the radius of the opening,

$$r^2 \frac{a+b}{ab} = 9\lambda$$

* *Collected Works*, Vol III p 429

CHAPTER VIII.

THE PROPAGATION OF LIGHT IN CRYSTALLINE MEDIA.

84. **The Ellipsoid of Plane Wave Propagation. Ellipsoid of Elasticity.** It has been shown in Chapter II, Art 12, that the velocity of propagation of a distortional wave in an isotropic medium is $\sqrt{\frac{n}{\rho}}$, where n is the resistance to distortion. If the medium is not isotropic, the coefficient n may depend on the direction of the displacement In that case, a plane wave may be propagated with different velocities according to the direction of the vibration. Fixing our mind on a wave-front parallel to a given plane, there must always be one direction for which n has a maximum value n_1, and one for which it has a minimum value n_2 There is therefore a maximum and minimum velocity of propagation $\sqrt{\frac{n_1}{\rho}}$ and $\sqrt{\frac{n_2}{\rho}}$ respectively for every plane, and two directions of vibration corresponding to these two velocities. If the displacement is neither in one nor the other of these two directions, it might be possible to imagine that the wave would be propagated with some intermediate velocity This is not, however, found to be the case, but the wave splits into two wave-fronts proceeding with the velocities $\sqrt{n_1/\rho}$ and $\sqrt{n_2/\rho}$, and these are the only velocities possible for a wave-front which is parallel to a given plane. If we change the direction of the plane, the velocities in general change also It may be proved that the two directions of displacement corresponding to the minimum and maximum coefficient of distortion, are always at right angles to each other If the direction of displacement be confined to one or other of these two directions, a plane wave may be propagated as a single plane wave. But in the general case, the displacement must be decomposed into its components in the two directions for which a single plane wave propagation is possible. The following construction connecting the velocities of the plane wave propagation in different directions, though originally suggested by

theoretical considerations, should at the present stage be considered simply as a representation of experimental facts.

Take some fixed point O (Fig 108) in the crystalline medium and imagine planes drawn through the point In each plane take two lines OP_1 and OQ_1 which coincide with the two possible directions of vibration. If v_1 be the velocity of the plane wave for a direction of vibration OP_1 and v_2 for a direction OQ_1, take the points P_1, P_2, Q_1, Q_2, so that $OP_1 = OP_2 = V/v_1$, $OQ_1 = OQ_2 = V/v_2$, V being the velocity of the wave propagation in vacuo* If the plane through O be altered in direction and the points P and Q marked off for each, it is found that these points lie on an ellipsoid, which may be called the ellipsoid of plane wave propagation It is also found that the points P and Q lie at the ends of the semiaxes of the central sections of this ellipsoid If the ellipsoid is given, we may therefore find the direction of vibration and the corresponding velocities of waves parallel to any given plane, by drawing the central section which is parallel to that plane The semiaxes of the ellipse in which the section cuts the ellipsoid give the directions of vibration, and the velocities are inversely proportional to the semiaxes.

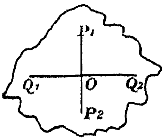

Fig. 108

Let the equation of the ellipsoid be

$$a^2x^2 + b^2y^2 + c^2z^2 = V^2 \qquad (1),$$

the quantities a, b, c, being in descending order To simplify the equation, take the unit of time to be the time which a wave in vacuo takes to traverse unit distance, so that we may write $V=1$ For $x=0$, we obtain the intersection of the ellipsoid with the plane of yz, which is an ellipse having $1/b$ and $1/c$ as semiaxes Hence a wave-front may be propagated in the direction of the axis of x either with a velocity b or with a velocity c, the direction of vibration in the former case being the axis of y, and in the latter the axis of z Similarly a and c are the velocities of propagation for a wave-front parallel to the plane xz, and a, b the velocities for a plane parallel to xy. Fig 109 illustrates how a plane wave separates into two, the directions of vibration in the two being at right angles to each other

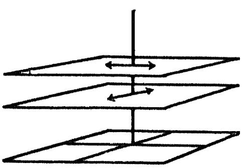
Fig 109

The problem of finding the velocities with which an inclined wave-front may be propagated, is one of geometry and may be solved

* We might also take $OP_1 = V/v_2$ and $OQ_1 = V/v_1$ and fit the observed phenomena equally well, but it is convenient to make our choice at once so as to fit in with the view that the direction of vibration is at right angles to the plane of polarization

as follows It is required to find the direction and magnitude of the principal axes OP and OQ of the intersection of the ellipsoid (1) by a central plane which is defined by the direction cosines l, m, n, of its normal ON, Fig 110

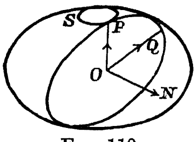

Fig. 110.

Let the length of the semiaxis OP be ρ Then a sphere of radius ρ and having its centre at O has one and only one point in common with the ellipse PQ, because the ellipse can only have *one* radius vector which has a length equal to that of one of its semiaxes The sphere intersects the ellipsoid in a curve which is called a sphero-conic, and this curve PS must touch the ellipse at P. If a cone be drawn having O as vertex, and passing through the sphero-conic PS, a similar reasoning shows that the cone cannot intersect, but must touch the plane of the section OPQ along the semiaxis OP

The equation to the cone is obtained by combining the equations of the ellipsoid

$$a^2x^2 + b^2y^2 + c^2z^2 = 1,$$

and the sphere

$$x^2 + y^2 + z^2 = \rho^2,$$

in such a way as to give the equation of a cone Multiplying the first equation by ρ^2 and subtracting, we obtain

$$(a^2\rho^2 - 1)\,x^2 + (b^2\rho^2 - 1)\,y^2 + (c^2\rho^2 - 1)\,z^2 = 0 \qquad \ldots.(2)$$

If the velocity of propagation of the plane wave is $v = 1/\rho$,

$$(v^2 - a^2)\,x^2 + (v^2 - b^2)\,y^2 + (v^2 - c^2)\,z^2 = 0$$

The direction cosines l, m, n, of the normal to the section must coincide with the direction cosines of the tangent plane of the cone, the line OP being the line of contact If x, y and z are the coordinates of P, we obtain, in the usual way,

$$D'l = (v^2 - a^2)\,x,$$
$$D'm = (v^2 - b^2)\,y,$$
$$D'n = (v^2 - c^2)\,z,$$

D' being determined by the condition that $l^2 + m^2 + n^2 = 1$.

We may introduce in place of x, y, z, the direction cosines α, β, γ of the vector OP The equations may then be written

$$\left.\begin{aligned} \frac{Dl}{(v^2 - a^2)} &= \alpha \\ \frac{Dm}{(v^2 - b^2)} &= \beta \\ \frac{Dn}{(v^2 - c^2)} &= \gamma \end{aligned}\right\} \qquad \ldots \qquad (3),$$

and D may now be determined by the condition that $\alpha^2+\beta^2+\gamma^2=1$ This gives

$$\frac{1}{D^2}=\frac{l^2}{(v^2-a^2)^2}+\frac{m^2}{(v^2-b^2)^2}+\frac{n^2}{(v^2-c^2)^2}$$

As the vector (α, β, γ) lies in a plane normal to the vector (l, m, n)

$$\alpha l+\beta m+\gamma n=0,$$

and hence we derive from equations (3)

$$\frac{l^2}{a^2-v^2}+\frac{m^2}{b^2-v^2}+\frac{n^2}{c^2-v^2}=0 \tag{4}$$

This equation allows the velocity v to be calculated It is of the second degree in v^2 and has therefore two positive roots, which from the nature of the problem must be real.

Having obtained v, we may, by means of equations (3), determine α, β, γ, the direction cosines of the directions of vibration, and we shall again have two solutions, one corresponding to OP and the other to OQ

The velocities a, b, c, are called the three principal velocities, and as, with the unit of time adopted, the velocity of light in vacuo is one, the reciprocal velocities $1/a$, $1/b$, $1/c$, measure quantities which in an isotropic medium would correspond to the refractive index These quantities are therefore called the principal refractive indices

Denoting these by μ_1, μ_2, μ_3, we may write the equation of the ellipsoid (1) in the form

$$\frac{x^2}{\mu_1^2}+\frac{y^2}{\mu_2^2}+\frac{z^2}{\mu_3^2}=1$$

The coefficients of elasticity, which measure the resistance to distortion in the principal planes, are proportional to a^2, b^2, c^2 respectively, so that these constants are intimately connected with the elastic properties of the medium The ellipsoid (1) has therefore been called the ellipsoid of elasticity (see also Art 103) In a homogeneous medium, $\mu_1=\mu_2=\mu_3$, and the ellipsoid of elasticity becomes a sphere, having a radius numerically equal to the refractive index

85. The Optic Axes. Every ellipsoid has two circular sections passing through that principal axis which is neither the largest nor the smallest It follows that there are two directions in which a plane wave-front has only a single velocity These two directions are called the "*optic axes.*" The radius of the circular section is $1/b$, and putting $\rho=1/b$ in the equation (2) of the cone, it reduces to

$$(a^2-b^2)\,x^2+(c^2-b^2)\,y^2=0$$

This is the equation of the two planes which contain the two circular

sections The two directions of single wave velocities are the normals to these planes, so that

$$l_1 = \pm\sqrt{\frac{a^2 - b^2}{a^2 - c^2}}, \quad m_1 = 0, \quad n_1 = \pm\sqrt{\frac{b^2 - c^2}{a^2 - c^2}} \qquad (5)$$

are the direction cosines of the optic axes

When a wave is propagated in the direction of one of the optic axes, the direction of vibration may be anywhere in the plane, as in the circular section, any direction may be considered to be an axis.

86. Uniaxal and Biaxal Crystals. In general a crystal has two optic axes and is then called "biaxal." If two of the principal axes are equal to each other, there is only one optic axis, which is the axis of x if $b = c$, and the axis of z when $a = b$ The crystal is then said to be a "uniaxal" crystal.

The ellipsoid of elasticity for uniaxal crystals when $a = b$, is the spheroid

$$a^2(x^2 + y^2) + c^2 z^2 = 1$$

and the equation (4) for determining the velocities of plane wave propagation becomes, writing θ for the angle between the optic axis and the normal,

$$\frac{\sin^2\theta}{a^2 - v^2} + \frac{\cos^2\theta}{c^2 - v^2} = 0,$$

or

$$v^2 = c^2\sin^2\theta + a^2\cos^2\theta.$$

Hence the velocity depends only on the angle which the normal to the wave-front makes with the axis of revolution of the spheroid

87 Wave-Surface The passage of waves through crystalline media is completely determined by the equation we have obtained for the propagation of plane waves, but it is often convenient to base our investigations on a surface which is the locus of equal optical distances measured from a point as centre. Such a surface, according to the definition of Art. 18, is called a "Wave-Surface" Its relation to the optical distance between parallel wave-fronts as deduced in the last paragraph may best be seen by applying Huygens' principle. Let a plane wave (Fig 111), WF, be propagated upwards and with points P_1, P_2, P_3, etc as centres construct the surfaces of equal optical distance, corresponding to unit time, *i e* the wave-surfaces ST and $S'T'$ The furthest distance to which the wave-front WF can have gone in the time is the tangent plane $W'F'$ to all the wave-surfaces, and by Huygens' principle, as explained in Art 16, this plane will actually be the position of the wave-front after unit time. The lines which join the centres of disturbance P_1, P_2, P_3, etc with the

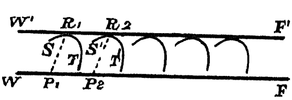

Fig 111.

points of contact R_1, R_2, R_3, etc of the wave-surfaces and wave-front, are the lines of shortest optical distance between WF and $W'F'$ These lines we have called the "rays" If the wave-surfaces are not spheres, the rays are not in general at right angles to the wave-fronts, and this is an important distinction between crystalline and isotropic media

If through any point P (Fig 112), we draw a number of plane wave-fronts, we may, from the results of the last article, construct the positions W_1F_1, W_2F_2, W_3F_3, etc of these wave-fronts after unit time Each wave-front must be a tangent plane to the wave-surface drawn with P as centre Hence the wave-surface is the envelope of all the plane wave-fronts Its equation may thus be obtained by a purely mathematical process

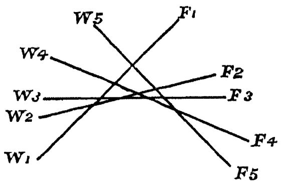

Fig 112

The equation to the wave-front is

$$lx + my + nz = v \qquad (6),$$

where v is the distance travelled in unit time, which itself is a function of l, m, n Any point Q of the wave-surface is a point of intersection of planes, differing from each other in direction by infinitely small quantities Hence a point x, y, z of the wave-surface must satisfy (6) and also the equation obtained by giving to l, m, n, v small increments dl, dm, dn, dv Subtracting one of these equations from the other it follows that

$$xdl + ydm + zdn = dv \qquad (7).$$

There are certain conditions to which the variations of the quantities l, m, n, v, are subject Thus

$$l^2 + m^2 + n^2 = 1$$

from which we derive

$$ldl + mdm + ndn = 0 \qquad (8),$$

and also as equation (4) continues to hold,

$$\frac{ldl}{v^2 - a^2} + \frac{mdm}{v^2 - b^2} + \frac{ndn}{v^2 - c^2} = Kvdv \qquad (9),$$

where
$$K = \frac{l^2}{(v^2 - a^2)^2} + \frac{m^2}{(v^2 - b^2)^2} + \frac{n^2}{(v^2 - c^2)^2}$$

As there are only two independent parameters of the plane, *i e.* l and m, it must be possible to express dn and dv in terms of dl and dm Two equations are sufficient for this purpose, and of the three equations (7), (8), (9) only two can be independent To express the condition that one of these equations may be obtained as a consequence of the two others, we multiply (8) by A and (9) by B, and add.

Equalizing the factors of dl, dm, dn, and dv, in this combined equation, with the corresponding factors in (7), we find

$$\left.\begin{aligned} x &= Al + B\frac{l}{(v^2-a^2)} \\ y &= Am + B\frac{m}{v^2-b^2} \\ z &= An + B\frac{n}{v^2-c^2} \end{aligned}\right\} \quad . \qquad (10),$$

$$BKv = 1 \qquad (11)$$

Multiplying (10) by l, m, n, respectively, and adding, we obtain, with the help of (6) and (4),

$$v = A \quad . \qquad (12)$$

Squaring and adding the equation (10), and writing r^2 for $x^2+y^2+z^2$, the term containing the product AB disappears in consequence of (4) and we obtain:

$$r^2 = A^2 + B^2K$$

With the help of (11) and (12) this gives B

$$B = v\,(r^2 - v^2) \quad \ldots. \qquad (13)$$

The first of equations (10) may now be written

$$x = \frac{B + A\,(v^2-a^2)}{v^2-a^2}\,l = \frac{r^2-a^2}{v^2-a^2}\,vl$$

$$\left.\begin{aligned} &\text{Hence} && \frac{x}{r^2-a^2} = \frac{vl}{v^2-a^2} = \frac{x-vl}{r^2-v^2} \\ &\text{Similarly} && \frac{y}{r^2-b^2} = \frac{vm}{v^2-b^2} = \frac{y-vm}{r^2-v^2}, \\ &&& \frac{z}{r^2-c^2} = \frac{vn}{v^2-c^2} = \frac{z-vn}{r^2-v^2} \end{aligned}\right\} \qquad (14)$$

Multiplying these equations by x, y, z, respectively and adding, the quantities l, m, n disappear owing to the relation (6) and we obtain

$$\frac{x^2}{r^2-a^2} + \frac{y^2}{r^2-b^2} + \frac{z^2}{r^2-c^2} = 1 \qquad (15).$$

This is one form of the equation of the wave-surface Another form is obtained by multiplying both sides of (15) by r^2, and then subtracting $x^2+y^2+z^2$ from the left-hand side, and r^2 from the right-hand side This leaves

$$\frac{a^2x^2}{r^2-a^2} + \frac{b^2y^2}{r^2-b^2} + \frac{c^2z^2}{r^2-c^2} = 0 \qquad (16),$$

a form which is usually more convenient than the former. We may also get rid of the denominators and write

$$(r^2-b^2)\,(r^2-c^2)\,a^2x^2 + (r^2-a^2)\,(r^2-c^2)\,b^2y^2 + (r^2-a^2)\,(r^2-b^2)\,c^2z^2 = 0.$$

Multiplying out and dividing by the common factor r^2 we find

$$r^2(a^2x^2 + b^2y^2 + c^2z^2) - a^2(b^2 + c^2)x^2 - b^2(c^2 + a^2)y^2 - c^2(a^2 + b^2)z^2 + a^2b^2c^2 = 0 \quad . \quad (17)$$

This is an equation of the fourth degree

88. Ray Velocity. The radius vector of the wave-surface, being the distance through which a disturbance may be considered to have travelled in unit time, measures the velocity along the ray Calling s the ray velocity, while v is the velocity of plane wave propagation, Fig 110 shows that if ζ be the inclination of the ray to the wave normal, $v = s\cos\zeta$, or if λ, μ, ν be the direction cosines of the ray,

$$v = s(l\lambda + m\mu + n\nu) \quad . \quad (18).$$

Equation (16) may serve to connect the direction λ, μ, ν with the velocity s, by writing s for r and using $x \; y \; z = \lambda \; \mu \; \nu$,

$$\frac{a^2\lambda^2}{s^2 - a^2} + \frac{b^2\mu^2}{s^2 - b^2} + \frac{c^2\nu^2}{s^2 - c^2} = 0 \quad . \quad . \quad . \quad (19)$$

Comparing this equation with (4) it is seen that one may be derived from the other by writing $1/s$ for v, and $1/a$, $1/b$, $1/c$, for a, b, c respectively This suggests the following construction for obtaining ray velocities similar to that which holds for the wave velocities Take the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

The lengths of the semiaxes of a plane central section of this ellipsoid then measure the two possible velocities of rays which are at right angles to the section. The proof of this proposition follows from the fact that we may obtain equation (19) to determine the lengths of the semiaxes by a reasoning identical with that by means of which (4) has been obtained in Art 84 The ellipsoid has been called the reciprocal ellipsoid Its semiaxes are equal to the principal velocities

89. Relations between rays and wave normals. The projections of the ray on the three coordinate axes are $s\lambda$, $s\mu$, $s\nu$, and we may substitute these values for x, y, z, in equations (14). We then obtain

$$\left.\begin{aligned} \frac{s\lambda}{s^2 - a^2} &= \frac{vl}{v^2 - a^2} \\ \frac{s\mu}{s^2 - b^2} &= \frac{vm}{v^2 - b^2} \\ \frac{s\nu}{s^2 - c^2} &= \frac{vn}{v^2 - c^2} \end{aligned}\right\} \quad . \quad . \quad . \quad (20).$$

Squaring and adding, we find

$$s^2\left[\frac{\lambda^2}{(s^2-a^2)^2}+\frac{\mu^2}{(s^2-b^2)^2}+\frac{\nu^2}{(s^2-c^2)^2}\right]$$
$$=v^2\left[\frac{l^2}{(v^2-a^2)^2}+\frac{m^2}{(v^2-b^2)^2}+\frac{n^2}{(v^2-c^2)^2}\right] \qquad (21)$$

We may also by a simple transformation of (20) obtain the three equations

$$vl-s\lambda=s\lambda\frac{v^2-s^2}{s^2-a^2},$$
$$vm-s\mu=s\mu\frac{v^2-s^2}{s^2-b^2},$$
$$vn-s\nu=s\nu\frac{v^2-s^2}{s^2-c^2},$$

and by squaring and adding, find

$$v^2+s^2-2vs\,(l\lambda+m\mu+n\nu)=s^2\,(v^2-s^2)^2\left[\frac{\lambda^2}{(s^2-a^2)^2}+\frac{\mu^2}{(s^2-b^2)^2}+\frac{\nu^2}{(s^2-c^2)^2}\right]$$

Introducing (18) this equation reduces to

$$\frac{1}{s^2-v^2}=s^2\left[\frac{\lambda^2}{(s^2-a^2)^2}+\frac{\mu^2}{(s^2-b^2)^2}+\frac{\nu^2}{(s^2-c^2)^2}\right] \qquad (22),$$

and hence from (21)

$$\frac{1}{s^2-v^2}=v^2\left[\frac{l^2}{(v^2-a^2)^2}+\frac{m^2}{(v^2-b^2)^2}+\frac{n^2}{(v^2-c^2)^2}\right] \quad . \qquad (23)$$

Equation (23) may be used to calculate the ray velocity from the velocity of the corresponding plane wave, while (22) is used when the ray velocity is given and the wave velocity is required

To determine the angle ζ included between s and v, we may use either

$$\cot^2\zeta=\frac{v^2}{s^2-v^2} \quad \text{or} \quad \operatorname{cosec}^2\zeta=\frac{s^2}{s^2-v^2}$$

The former gives, in terms of the quantities defining the plane wave propagation,

$$\cot^2\zeta=v^4\left[\frac{l^2}{(v^2-a^2)^2}+\frac{m^2}{(v^2-b^2)^2}+\frac{n^2}{(v^2-c^2)^2}\right],$$

and the latter in terms of the quantities defining the ray propagation

$$\operatorname{cosec}^2\zeta=s^4\left[\frac{\lambda^2}{(s^2-a^2)^2}+\frac{\mu^2}{(s^2-b^2)^2}+\frac{\nu^2}{(s^2-c^2)^2}\right]$$

The plane containing the direction of vibration OP, Fig 113, and the wave normal ON contains other vectors which are related to the wave propagation To express analytically the condition that three vectors should be in the same plane, we make use of the well-known relation between the direction cosines l, m, n, and ξ, η, ζ, of two lines which are at right angles to each other This condition is

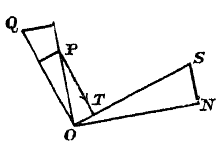

Fig 113

$$l\xi+m\eta+n\zeta=0$$

If ξ, η, ζ, is also at right angles to the vector λ, μ, ν,

$$\lambda\xi + \mu\eta + \nu\zeta = 0,$$

also

$$\xi^2 + \eta^2 + \zeta^2 = 1$$

Solving these equations, we find

$$\xi = m\nu - n\mu,$$
$$\eta = n\lambda - l\nu,$$
$$\zeta = l\mu - m\lambda$$

A third vector α, β, γ will be in the same plane with l, m, n and λ, μ, ν, if it is at right angles to ξ, η, ζ Hence

$$\alpha(m\nu - n\mu) + \beta(n\lambda - l\nu) + \gamma(l\mu - m\lambda) = 0$$

is the required condition If there is a linear relation between the three vectors, such that

$$\alpha = C\lambda + Fl, \quad \beta = C\mu + Fm, \quad \gamma = C\nu + Fn,$$

the condition is obviously satisfied Giving to the direction cosines their previous meaning, we have according to (3)

$$\left.\begin{aligned} Dl &= (v^2 - a^2)\,\alpha \\ Dm &= (v^2 - b^2)\,\beta \\ Dn &= (v^2 - c^2)\,\gamma \end{aligned}\right\} \qquad (24),$$

where D according to (23) is equal to $v\sqrt{s^2 - v^2}$

Also with the help of (20) and (24)

$$\left.\begin{aligned} E\lambda &= (s^2 - a^2)\,\alpha \\ E\mu &= (s^2 - b^2)\,\beta \\ E\nu &= (s^2 - c^2)\,\gamma \end{aligned}\right\} \qquad (25),$$

where E is written for Ds/v, which is equal to $s\sqrt{s^2 - v^2}$

Combining (24) and (25) we obtain the following linear relations·

$$E\lambda - Dl = (s^2 - v^2)\,\alpha,$$
$$E\mu - Dm = (s^2 - v^2)\,\beta,$$
$$E\nu - Dn = (s^2 - v^2)\,\gamma,$$

which proves that the three vectors are coplanar The "ray" therefore lies in the plane which contains the wave normal and the displacement In Fig 113 the wave normal and ray are indicated by the direction of the lines ON and OS We may now prove that the normal PT to the ellipsoid of elasticity at the point at which the direction of the displacement intersects it, lies in the same plane At the end of the radius vector OP draw the tangent PK to the ellipse of intersection If OP is a semiaxis, PK is at right angles to OP, and also to PT, the normal to the ellipsoid Hence PT and PO are in a plane which is at right angles to PK, and hence also at right angles to OP', the second semiaxis of the ellipse OP' being the normal to the plane containing

PT and PO, every line at right angles to OP' must lie in the same plane The normal to the section is such a line, hence PO, PT and ON are in the same plane

The direction cosines of the normal PT are proportional to a^2x, b^2y, c^2z, if x, y, z are the coordinates of P, and hence also proportional to $a^2\alpha$, $b^2\beta$, $c^2\gamma$

The ray OS is at right angles to that plane section of the reciprocal ellipsoid which has the ray velocity s as one of its semiaxes If we proceed exactly as in Art 84 to find the direction cosines of the two semiaxes of the ellipse which is at right angles to λ, μ, ν, we find by equations analogous to (3) that they are proportional to

$$a^2\lambda/(s^2-a^2),\quad b^2\mu/(s^2-b^2),\quad c^2\nu/(s^2-c^2),$$

and therefore by comparison with (20) proportional to $a^2\alpha$, $b^2\beta$, $c^2\gamma$

This proves that the semiaxis in question is parallel to PT Let OQ (Fig 113) be that semiaxis The normal to its tangent plane has direction cosines proportional to $\frac{\alpha'}{a^2}$, $\frac{\beta'}{b^2}$, $\frac{\gamma'}{c^2}$, if α', β', γ' fix the direction of OQ From the ratio $\alpha' \ \beta' \ \gamma'$ which has just been found, it follows at once that OP is the normal to the tangent plane at Q

The second semiaxis of the section of the ellipsoid of elasticity passing through OP is coincident in direction with the second semiaxis of the section of the reciprocal ellipsoid, which passes through OQ, because in both cases the semiaxis must be at right angles to the plane of the figure

90 The direction of displacement It has been proved that the vectors OP, OS, and ON (Fig 113) are in the same plane, OP indicating the direction of the displacement As NS is in the wave-front, the vibration takes place in the direction NS, which is the projection of the ray on the wave-front. The direction of vibration cannot be observed, and the above statement involves therefore something that is theoretical and based on a particular assumption as to the nature of light That assumption has been introduced by the manner in which the construction of Art 84 has been carried out, as already explained in the footnote to that article If we wish to confine our statements to facts capable of experimental verification, we ought to say that the plane of polarization is at right angles to the projection of the ray on the wave-front. The two ways of expressing the facts are identical, if the direction of vibration is at right angles to the plane of polarization

The planes of polarization may be obtained in a simple manner from the direction of the optic axes The axes of an ellipse are the bisectors of the angles formed by any two equal diameters, and

the planes which are normal to the optic axes intersect the ellipsoid of elasticity in a circle of radius $1/b$. Hence if ON (Fig 114) represents the normal to the wave-front, and $PP'Q'Q$ the elliptical intersection of the wave-front and the ellipsoid, two lines OP, OP' which are at right angles to the optic axes OH and OH' are of equal length $1/b$ A plane through ON and OH intersects the ellipse in a line OQ at right angles to OP, and a plane through the normal and second optic axis intersects the ellipse along OQ' at right angles to OP' Hence OQ and OQ' are equal radii of the ellipse It follows that the planes of polarization are the two bisectors of the planes which pass through the normal and two optic axes respectively

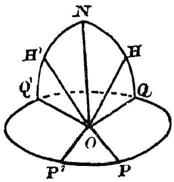

Fig 114

The two plane waves propagated in the same direction have their planes of polarization at right angles to each other, but the two planes of polarization belonging to a given *ray* are not at right angles to each other unless the ray coincides with the wave normal To prove this, we take two directions of vibration α, β, γ, and α_1, β_1, γ_1, which correspond to the same value of λ, μ, ν, the ray velocities being s and s_1, and the corresponding wave velocities v and v_1 According to (25) we have

$$E\lambda = (s^2 - a^2)\,\alpha\,,\quad E\mu = (s^2 - b^2)\,\beta\,,\quad E\nu = (s^2 - c^2)\,\gamma,$$

$$E_1\lambda = (s_1^2 - a^2)\,\alpha_1\,,\quad E_1\mu = (s_1^2 - b^2)\,\beta_1\,,\quad E_1\nu = (s_1^2 - c^2)\,\gamma_1,$$

where $\quad E = s\sqrt{s^2 - v^2}$ and $E_1 = s_1\sqrt{s_1^2 - v_1^2}$

The cosine of the angle ω between the directions of vibration is

$$\cos\omega = \alpha\alpha_1 + \beta\beta_1 + \gamma\gamma_1,$$

and after substitution, the right-hand side is found to be equal to

$$\frac{EE_1}{s_1^2 - s^2}\left[\left(\frac{\lambda^2}{s^2 - a^2} + \frac{\mu^2}{s^2 - b^2} + \frac{\nu^2}{s^2 - c^2}\right) - \left(\frac{\lambda^2}{s_1^2 - a^2} + \frac{\mu^2}{s_1^2 - b^2} + \frac{\nu^2}{s_1^2 - c^2}\right)\right]$$

With the help of (19) this becomes

$$\frac{EE_1}{s^2 s_1^2} = \frac{\sqrt{(s^2 - v^2)(s_1^2 - v_1^2)}}{ss_1}$$

Hence if ϕ and ϕ_1 are the angles included between the common direction of the ray and the two wave normals,

$$\cos\omega = \sin\phi\,\sin\phi_1 \qquad (26)$$

In order that the two directions of vibration should be at right angles to each other, it is therefore necessary that the ray should coincide with one of the wave normals

91. Shape of the wave-surface. We may now form an idea of the general shape of the wave-surface If in (17) we put successively $z = 0$, $y = 0$, $x = 0$, we obtain the intersections of the wave-surface with the coordinate planes.

The intersection with the plane of xy is

$$(x^2+y^2)(a^2x^2+b^2y^2)-a^2(b^2+c^2)x^2-b^2(c^2+a^2)y^2+a^2b^2c^2=0,$$

or

$$(x^2+y^2-c^2)(a^2x^2+b^2y^2-a^2b^2)=0$$

This is separately satisfied by

$$x^2+y^2-c^2=0,$$

and by

$$a^2x^2+b^2y^2-a^2b^2=0$$

The curve of intersection breaks up therefore into a circle radius c and ellipses of semiaxes a and b The circle lies inside ellipse and does not intersect it, because we have assumed c to smaller than both a and b

The intersection with the plane of yz is similarly found to b circle of radius a and an ellipse of semiaxes b and c Here the ci lies completely outside the ellipse

The intersection of the wave-surface with the plane xz splits into a circle

$$x^2+z^2=b^2,$$

and an ellipse

$$c^2z^2+a^2x^2-a^2c^2=0$$

The circle and ellipse in this plane intersect (Fig 115), the points of intersection being given by

$$\left.\begin{aligned} x &= \pm c\sqrt{\frac{a^2-b^2}{a^2-c^2}} \\ z &= \pm a\sqrt{\frac{b^2-c^2}{a^2-c^2}} \end{aligned}\right\} \quad \ldots\ldots (27$$

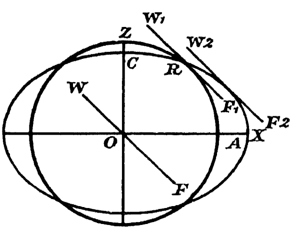

Fig 115

Fig 116 represents in perspective intersections of the wave-surface with three coordinate planes, one quadrant being drawn. The letters attached to curves of intersection represent the lengths of the semiaxes, thus means a circle of radius c. The c plete wave-surface consists of sheets, an inner sheet, and an o sheet which meets the inner shee four points, one lying in each of four quadrants of the plane xz. coordinates of these four points those given above (27)

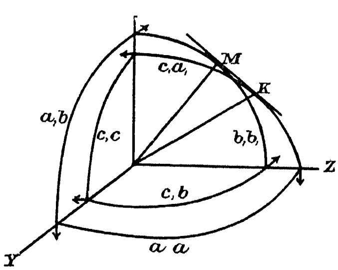

Fig 116

The directions of vibration indicated in the figure by arr Each ray which lies in a prin plane coincides with the normal to the wave-surface at one of

points of intersection with it It therefore coincides with one of the normals of the wave-fronts which may belong to it Hence according to (26) a ray which lies in one of the principal planes, has its two directions of vibration at right angles to each other On the elliptical intersections of the wave-front and the principal planes, the direction of vibration, being as proved above the projection of the ray on the tangent plane, must lie in the principal plane It follows that in the circular intersections of the wave-surface and the principal planes, the direction of vibration is at right angles to the plane

92. The axes of single ray velocity and of single wave velocity. In general, any straight line drawn in one direction from the origin, intersects the wave-surface in two points, one on each of the two sheets The intercepts between the origin and the points of intersection measure the optical lengths There are therefore in each direction in general, two different optical lengths according to the direction of vibration But the four directions OR (Fig 115) form an exception, for there is only one point of intersection for each of them, and therefore only one optical length If a wave is propagated through the crystal, and the ray happens to lie along the direction OR, the difference in optical length between two points along the ray is independent of the direction of the displacements Adopting the definition of "ray velocity" given in Art 88 we may call these directions the axes of "single ray velocity" They are not coincident with the optic axes The latter indicate the direction of single *wave* velocity, the wave being considered to be plane, and normal to one of the optic axes. Remembering that we may obtain the position of a plane wave-front WF, Fig 115, after unit time, if we construct the wave-surface and draw the tangent planes to that surface parallel to WF, we see from the figure that in general (as we know already), there are two tangent planes W_1F_1, and W_2F_2, which are parallel to each other, and to WF The lengths of the perpendiculars from O on these planes, measure the wave velocities If there is a direction in which there is only one wave velocity, the two tangent planes normal to that direction must coincide There is indeed one tangent plane MK (Fig 116) in each quadrant, which touches both sheets of the wave-surface simultaneously. Symmetry shows that these tangent planes are parallel to OY, and as they must touch the circle of radius b in the plane of the figure at a point M, OM must be the direction of one of the optic axes Combining (27) and (5) we find for the cosine of the angle between the optic axes and the axes of single ray velocities,

$$\frac{b^2+ca}{b(a+c)}.$$

If b is equal to either a or c, or if a and c differ but little from each other, the angle is small, and may sometimes be neglected To form an idea of the error committed in this way, write $c = b(1 - \delta)$ and $a = b(1 + \epsilon)$, where δ and ϵ are small We obtain for the cosine of the angle included between the two sets of axes to the second approximation

$$1 - \tfrac{1}{2}\delta\epsilon \tag{28}$$

For mica the angle is about 40′ Even in the case of asparagin, a crystal in which the optic axes are nearly 90° apart, the angle between the optic axis and the axis of single ray velocity is less than 2°

93 Peculiarity of single wave propagation. In general, one ray belongs to each plane wave which is propagated through a crystalline medium, and the radii drawn from O (Fig 115) to the points of contact of the parallel planes W_1F_1 and W_2F_2 with the wave-surface are the rays belonging to the two waves propagated parallel to WF When the wave normal coincides with the optic axis, there is only one velocity as we have seen, and inspection of Figure 116 shows that the two rays OM and OK belong to this same wave But the wave-front WF touches the wave-surface not only at the two points M and K, but along the complete circumference of a circle drawn with MK as diameter To prove this, we must turn back to equations (14) which determine the points of contact x, y, z, of a plane defined by the direction cosines of its normal (l, m, n) and the wave-surface To suit our present problem, we must put $m = 0$, $v = b$ The second equation becomes indeterminate and may be satisfied for any value of y, by a proper choice of the two indefinitely small quantities m and $v - b$ The first and third equations are therefore the only ones we need consider Multiplying the first by l, and the third by n we obtain

$$\begin{aligned}\frac{xl}{r^2 - a^2} + \frac{zn}{r^2 - c^2} &= v\left(\frac{l^2}{v^2 - a^2} + \frac{n^2}{v^2 - c^2}\right)\\ &= v\left(\frac{l^2}{b^2 - a^2} + \frac{n^2}{b^2 - c^2}\right)\\ &= 0\end{aligned}$$

The last step follows on substituting for l, n the direction cosines of the optic axis Hence

$$lx(r^2 - c^2) + nz(r^2 - a^2) = 0 \tag{29}.$$

Also

$$xl + zn = b \tag{30},$$

which may be similarly obtained from (14) The equations (29) and (30) combined give

$$b(x^2 + y^2 + z^2) = lc^2x + na^2z \quad \ldots \tag{31},$$

which is the equation of a sphere passing through the origin of coordinates Equation (30) represents the tangent plane to the wave-surface at the point M

The point of contact of the tangent plane and the wave-surface is therefore the same as the intersection of the tangent plane and the sphere (31) It follows that the line of contact is a circle The rays which join O to a point on the circle form a cone, the equation to which may be obtained by multiplying (30) and (31) together, *i e*

$$b^2(x^2+y^2+z^2) = (lc^2x + na^2z)(lx+nz)$$

It is a remarkable fact that we may have a plane wave propagation, such that the condition of minimum optical length from a point P is satisfied not only for one direction, but for an indefinite number of directions lying on a cone, CC' (Fig 117) For any point T which lies either inside or outside the circle forming the base of the cone, the optical length is greater It should, however, be noticed that if the wave is plane polarized, there is only one ray The distinction between this case and the general one is therefore that while in general the vibration of an unpolarized wave-front may be decomposed into two, for either of which there is a definite wave velocity and a corresponding ray, the vibrations do not in this special case divide themselves into two components, but to each direction of vibration belongs a different ray, all these rays lying on a cone of the second degree.

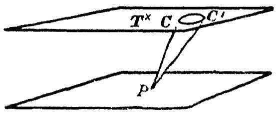

Fig 117

94. Peculiarity of a single ray propagation. We may obtain results analogous to the preceding ones, if we try to find the directions of the normals to the tangent planes at the conical point where the direction of single ray velocity cuts the wave-surface We make use for this purpose of equations (20) With the same notation as before $x = vl$, $z = vn$, are the coordinates of the foot of the perpendicular from the origin to any one of the tangent planes The ray velocity being constant and equal to b while $\mu = 0$, the second equation is indeterminate and the first and third become

$$\frac{b\lambda}{b^2-a^2} = \frac{x}{r^2-a^2},$$

$$\frac{b\nu}{b^2-c^2} = \frac{z}{r^2-c^2}.$$

With the help of (27) considering that x and y in that equation are proportional to λ, μ,

$$\frac{a^2\lambda x}{r^2-a^2} + \frac{c^2\nu z}{r^2-c^2} = 0,$$

and
$$a^2\lambda x + c^2\nu z = \frac{a^2c^2}{b}.$$

The first equation, with the help of the second, becomes

$$r^2 = b(\lambda x + \nu z)$$

Hence the locus of the foot of the perpendicular from O to the tangent planes at R is the circle formed by the intersection of a sphere with a plane The equation of the plane shows that it is parallel to OY and touches at R the ellipse ARC, Fig 115

95 Wave-surface in uniaxal crystals. The wave-surface in uniaxal crystals takes the shape already indicated by Huygens If $b = c$ equation (17) reduces to

$$(r^2 - c^2)(a^2x^2 + c^2y^2 + c^2z^2 - a^2c^2) = 0$$

The surface splits up therefore into the sphere of radius c and the spheroid

$$\frac{x^2}{c^2} + \frac{y^2 + z^2}{a^2} = 1$$

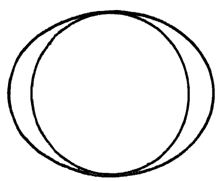

Fig 118

Similarly, if $a = b$, the equation of the wave-surface splits up into a sphere of radius a and into the ellipsoid

$$\frac{x^2 + y^2}{c^2} + \frac{z^2}{a^2} = 1.$$

Figs 118 and 119 represent the two cases, the optic axes being in the first case the axis of X and in the second case the axis of Z The positions of the axes are determined if we take a, b, c to be always in descending order, but if we drop that supposition, we may take the optic axis of uniaxal crystals to be at our choice either in the direction of X or in the direction of Z. The spheroid is formed by the revolution of the ellipse and circle round the optic axis. The two types of wave surfaces, one having an oblate and the other a prolate spheroid, according as the generating ellipse is made to rotate about its shorter or longer diameter, are illustrated by the case of Iceland Spar and Quartz The term positive and negative crystals, as applied to crystals similar to Quartz and Iceland spar respectively, is confusing and should be avoided We may speak instead of crystals which are optically prolate, or optically oblate, and in a discussion relating to optical properties only, where no confusion is possible, we may call them shortly prolate and oblate crystals.

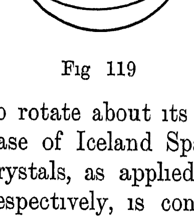

Fig 119

96. Refraction at the Surface of Uniaxal Crystals. The refracted waves may, in crystalline media, be constructed exactly as in

isotropic bodies, but as the wave-surface consists of two sheets, there are in general two refracted rays. In uniaxal crystals, one sheet of the wave-surface is always a sphere, and hence one of the rays follows the ordinary laws of refraction This ray is called the ordinary ray, and the ratio of the sines of the two angles of direction is called the ordinary refractive index

As regards the second ray, it will not in general follow any simple law, and may or may not lie in the plane of incidence If the optic axis is inclined to the plane of incidence, the point of contact of the spheroid with the sphere does not lie in that plane, Fig 120, and if the plane BT is drawn in the usual way at right angles to the plane of incidence, to touch the spheroidal sheet of the wave-surface, the ray being the line joining O to the point of contact does not lie in that plane The ordinary ray OK is found in the usual way and always remains in the plane of incidence.

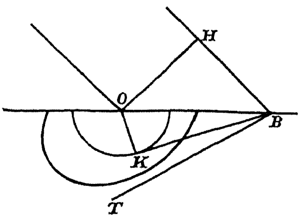

Fig 120

In Fig. 120 HB is drawn parallel to the incident ray and at such a distance from it that HB is unit length The scale of the wave-surface is such that it represents the locus of the disturbance spread out from O in unit time, which for the present purpose has been chosen to be such that the velocity of light in vacuo is unity BT represents the trace of the refracted wave which touches the ellipsoid at a point outside the plane of incidence The refracted ray does not in this case lie in that plane

It is not necessary to obtain the general equation, giving the direction of the refracted ray, and we may treat a few special cases separately.

(*a*) *The optic axis of the crystal at right angles to the plane of incidence* The trace of the wave-surface on the plane of incidence is in this case a circle, and the refracted ray may by symmetry be seen to lie in the plane of incidence Hence the rays follow the ordinary law of refraction In oblate crystals the outer circle of radius a belongs to the extraordinary ray, and its angle of refraction is greater. The reverse holds for prolate crystals For oblate crystals, the ratio of the sines for the extraordinary ray is with the unit time chosen $1/a$. Calling this μ_e we may write for the equation to the wave-surface

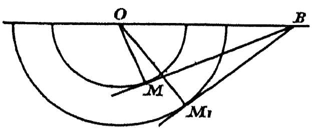

Fig 121

$$\mu_0^2 x^2 + \mu_e^2 (y^2 + z^2) = 1,$$

where μ_0 is the refractive index of the ordinary ray. The extraordinary

refractive index μ_e has obtained its name and significance from the optic behaviour of the extraordinary ray in the general case we are now considering In the case of prolate crystals, the equation of the wave-surface in terms of the principal refractive indices becomes

$$\mu_e^2(x^2+y^2)+\mu_0^2 z^2=1,$$

the axis of z being now the optic axis

(*b*) *The optic axis is in the surface and plane of incidence* The refracted rays are both in the plane of incidence From the projective properties of the ellipse,

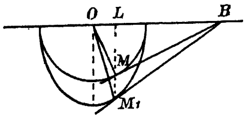

Fig 122

$$\frac{LM}{LM_1}=\frac{c}{a},$$

and if the angles of refraction of the *wave normals* are ϕ and ϕ_1,

$$\frac{\tan\phi}{\tan\phi_1}=\frac{\tan OBM}{\tan OBM_1}=\frac{LM}{LM_1}=\frac{\mu_e}{\mu_0},$$

an equation which holds for both prolate and oblate crystals

If the angles of the refraction of the *rays*, OML and OM_1L, be denoted by r and r_1, we obtain similarly

$$\frac{\tan r}{\tan r_1}=\frac{\mu_0}{\mu_e}.$$

(*c*) *The optic axis is perpendicular to the refracting surface* If P_1AQ_1 be the trace of the ellipsoid on the plane of incidence, and if we construct a circle with $P_1Q_1=2a$ as diameter, we have, writing ϕ_1 for the angle LBM, which is the angle of refraction of the extraordinary wave,

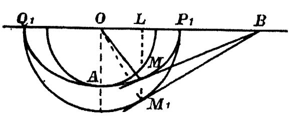

Fig 123

$$\frac{\tan\phi_1}{\tan LBM_1}=\frac{LM}{LM_1}=\frac{c}{a}.$$

But $$\sin LBM_1=\frac{OM_1}{OB},$$

and if, for OM_1, we may put its value a, and for OB, $1/\sin i$,

$$\tan\phi_1=\frac{c\sin i}{\sqrt{1-a^2\sin^2 i}}$$

or introducing $$\mu_0=1/c,\ \mu_e=1/a,$$

$$\tan\phi_1=\frac{\mu_e\sin i}{\mu_0\sqrt{\mu_e^2-\sin^2 i}}$$

And similarly if r_1 is the angle of refraction of the extraordinary ray,

$$\tan r_1=\frac{\mu_0\sin i}{\mu_0\sqrt{\mu_e^2-\sin^2 i}}$$

(*d*) *The incident wave-front is parallel to the surface* Let the plane of the paper (Fig 124) contain the optic axis The refracted extraordinary ray lies along OM, where M is the point of contact of the spheroidal portion of the wave-surface with a plane drawn parallel to the surface The ordinary ray coincides, of course, with the normal ON To determine the angle between the two rays, which is also the angle of refraction of the extraordinary ray, we must obtain an expression for the angle between the radius vector OM of an ellipse, and the normal to its tangent at M If θ be the angle between the optic axis and the surface which is equal to the angle between ON and OH, the major axis of the ellipse, and γ be the angle between OM and OH, we have for the properties of the ellipse,

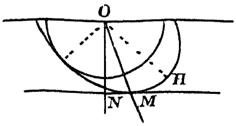

Fig 124

$$\tan\theta = \frac{a^2}{c^2}\tan\gamma$$

Hence if r is the angle of refraction of the extraordinary ray,

$$\begin{aligned}\tan r = \tan(\theta-\gamma) &= \frac{\tan\theta - \tan\gamma}{1+\tan\theta\tan\gamma}\\ &= \frac{(a^2-c^2)\tan\theta}{a^2+c^2\tan^2\theta}\\ &= \frac{\mu_0^2-\mu_e^2}{\mu_0^2\cot\theta+\mu_e^2\tan\theta}\end{aligned}$$

97. Direction of vibration in uniaxal crystals The rule that the direction of vibration is in the direction of the projection of the ray on the wave-front shows at once that on the spheroidal portion of the wave-front, the direction of vibration must be in a plane containing the optic axis As the condition (Art 90) under which the two vibrations along the same ray are at right angles to each other, always holds in uniaxal crystals, we may say that the ordinary ray is always polarized in a principal plane, and the extraordinary ray at right angles to that plane

98. Refraction through a crystal of Iceland Spar. A crystal of Iceland Spar is a rhomb (Fig 125) The parallelograms forming its six faces have sides which include angles of 102° and 78° respectively The faces are inclined to each other at angles of 105° and 75° There are two opposite corners A and B at which the three edges all form obtuse angles of 102° with each other The optic axis is parallel to the line drawn through one of these corners A, and equally inclined to the three faces

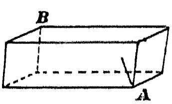

Fig 125

Double refraction may easily be exhibited by placing such a rhomb on a white sheet of paper on which a sharp mark is drawn When this mark is looked at from above through the crystal, it appears double, and if the crystal be turned round, one image seems to revolve round the other Let O, Fig 126, be the mark, the images of which are observed To trace the image formed by the extraordinary ray, construct a wave-surface to such a scale that the spheroid touches the face DC If T is the point of contact, a ray OT is refracted outwards along the normal TM, because at T the tangent plane to the wave-surface and the surface are coincident The refraction is therefore the same at that point as for a wave incident normally

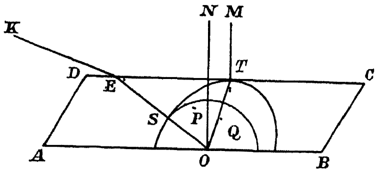

Fig 126

A ray OS parallel to the optic axis intersects the face at a point E, and is refracted along some direction EK. Disregarding aberrations, the intersection Q of KE and TM gives the extraordinary image As there can be no distinction between an ordinary and an extraordinary ray along the optic axis, the ordinary image P is obtained by the intersection of the same line EK with the normal ON, on which the ordinary image must lie The figure shows that this ordinary image lies nearer to the surface than the extraordinary one, and if the crystal be turned round the point O, the image Q travels in a circle round P The vertical plane containing P and Q contains also the optic axis, and the ordinary image is therefore polarized in the plane which passes through the two images, the extraordinary image being polarized at right angles to it

99. Nicol's Prism. A Nicol's prism, or, as it ought to be more appropriately called, a Nicol's rhomb, is one of the most useful appliances we have for the study of polarization Let Fig 127 represent the section of a long rhomb of Iceland Spar, passing through the optic axis, and LL' an oblique section through it If the rhomb be cut along this section and then recemented together by means of a thin layer of Canada balsam, only rays polarized at right angles to the principal plane are transmitted through it, if the inclination of the section LL' has been properly chosen An unpolarized ray is refracted at the surface, and separated into two, the extraordinary ray being bent less away from the original direction The ordinary ray falls therefore more obliquely on the surface of

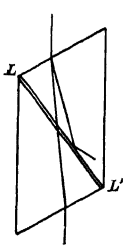

Fig 127

separation LL' The velocity of light in Canada balsam being intermediate between that of the two sets of waves in Iceland Spar, the inclination of LL' may be adjusted so that the ordinary ray is totally reflected, while the extraordinary ray passes through the combination Fig 124 shows in perspective how the plane of division is cut through the rhomb When the end face $ABCD$ of the rhomb is a parallelogram and parallel to one of the cleavage planes, the inclination of the section must be such that the side BB' of the rhomb is about 3 7 times as long as one of the sides of the end faces It is difficult to secure crystals of Iceland Spar which are sufficiently long to give, under these conditions, a beam of such cross section as is generally required in optical work The angular space through which the Nicol prism is effective in polarizing light is determined by the fact that if the incidence on the face LL' is too oblique, the extraordinary ray is totally reflected as well as the ordinary ray, and if not oblique enough, the ordinary ray can pass through The field of view containing the angular space thus limited when the prism is cut according to the above directions, is about 30° If it is not necessary to have so wide a field of view, shorter lengths of crystals can be used by cutting the end face $ABCD$, so as to be more nearly perpendicular to the length Sometimes that face is even inclined the other way. A field of view of 25° may thus be secured with a ratio of length to breadth of 2 to 5 Artificial faces at the end have, however, the disadvantage of deteriorating more quickly than cleavage planes

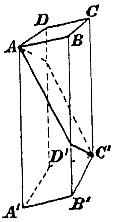

Fig 128

Foucault constructed a rhomb in which a small thickness of air is introduced in place of the Canada balsam. The prism need then be barely longer than broad, but the field of view is reduced to 7°

100. Double Image Prisms. It is sometimes convenient to have two images of a source near together, achromatic as far as possible, and polarized perpendicularly to each other An ordinary prism made of Iceland spar or quartz cannot be used on account of the colour dispersion, but if a prism of quartz be achromatised by means of a prism of another material, the desired result may be obtained If glass is chosen for the material of the second prism, the achromatism is only complete for one of the images, but for many purposes it is sufficiently perfect for the second image also The purpose is better obtained by prisms, like that of Rochon, in which the same material is used for both prisms, but turned differently with respect to the optic axis In Rochon's

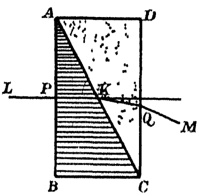

Fig 129

arrangement the optic axis of the first prism ABC (Fig 129) is parallel to the normal BC, this being indicated in the figure by the direction of the shading A ray LP incident normally is propagated without change of direction The axis of the second prism ACD is at right angles to the plane of the figure, and double refraction takes place at K, one ray being propagated in the normal direction as before, but the extraordinary ray being refracted along KQ and, on passing out of the prism, along QM The achromatism is complete for the image formed by the ordinary ray, and nearly complete for the other In the prism of Wollaston (Fig 130), the axis of the first prism is parallel to AB and that of the second at right angles to the plane of the figure; the path of the rays is indicated in the figure

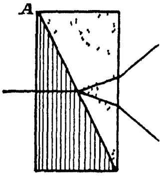

Fig 130

101 Principal Refractive Indices in biaxal crystals If refraction takes place at the surface of a biaxal crystal, and the plane of incidence is one of the principal planes (*e g* the plane of YZ), both rays lie in the plane of incidence A plane wave-front incident at O must, after refraction, touch a circle of radius a, and an ellipse of semiaxes b and c which form the intersection of the wave-surface with the plane of YZ One of the rays follows the ordinary law of refraction, while the angle of refraction of the other ray may be obtained as in case (c), Art 96 The refractive index of the rays belonging to the circular section is $1/a$, similarly for planes of incidence coincident with the planes of XZ and of YZ, we should have always one ray following the ordinary law, the corresponding refractive indices being $1/b$ and $1/c$ These three quantities are therefore called the principal refractive indices. Denoting these by μ_1, μ_2, μ_3, we may express all quantities relating to the wave surface in terms of them Thus for the direction cosines of the optic axes, we may put, using (5),

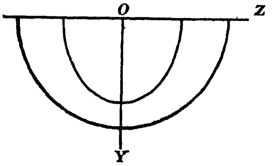

Fig 131

$$l_1 = \pm \frac{\mu_3}{\mu_2}\sqrt{\frac{\mu_2^2 - \mu_1^2}{\mu_3^2 - \mu_1^2}}, \quad m_1 = 0; \quad n_1 = \pm \frac{\mu_1}{\mu_2}\sqrt{\frac{\mu_3^2 - \mu_2^2}{\mu_3^2 - \mu_1^2}},$$

and for the direction cosines of the rays of single ray velocity, using (27),

$$\lambda = \pm \sqrt{\frac{\mu_2^2 - \mu_3^2}{\mu_3^2 - \mu_1^2}}, \quad \mu = 0, \quad \nu = \pm \sqrt{\frac{\mu_1^2 - \mu_2^2}{\mu_3^2 - \mu_1^2}}.$$

102 Conical Refraction. Two cases of refraction in biaxal crystals have a special interest If a wave-front WF is incident on

a plate cut out of the crystal at an angle such that the refracted wave-front *HKLM* is normal to an optic axis, the ray *PD* may, according to the direction of vibration, be refracted along any direction lying on the surface of the cone investigated in Art 93, the cone intersecting the wave-front inside the crystal in a circle (Fig 132)

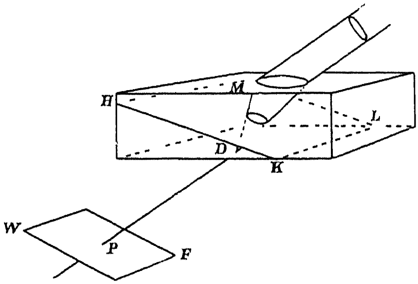

Fig 132

If the wave-front *WF* contains a number of coincident rays, having their planes of polarization symmetrically distributed in all directions, the refracted rays form the surface of a cone of the second degree which becomes a cylinder on emergence at the upper surface This interesting result was first deduced theoretically by Sir Wm Hamilton, from the shape of the wave-surface, and was afterwards experimentally verified by Lloyd To illustrate it experimentally, we may take a plate (Fig 133), cut so that its face is equally inclined to both axes An opaque plate *PQ* with a small aperture *O*, covers the side on which the light is incident A second plate *P′Q′* transmits light through a small hole at *O′*, which, if properly illuminated, may be considered to act as a source of light. If now *PQ* be moved along the face of the crystal, a direction *O′O* may be found such that if the original light is unpolarized, the ray *O′O* splits into a conical pencil, which may be observed after emergence at *AB*. This phenomenon is called "internal conical refraction" to distinguish it from another similar effect which takes place when a ray travels along an axis of single ray velocities

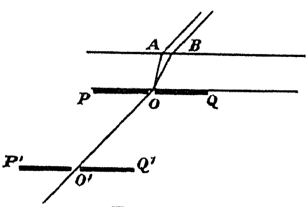

Fig 133

We may always follow the refraction of a ray belonging to a certain wave-surface and incident internally on the face of a crystal by considering it to be part of a parallel beam The wave-front belonging to this parallel beam would be the plane which touches

the incident wave-surface at the point of incidence If now a ray HR (Fig 134) travels inside a crystal along the axis of single ray velocity, there is an infinite number of tangent planes to the wave-surface at the point R, the normals of the tangent planes forming a cone HKL with a circular section at right angles to HR

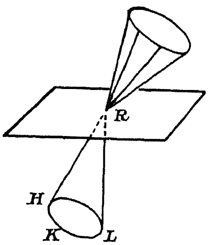

Fig 134

To each of these normals corresponds a separate ray on emergence and each ray has its own plane of polarization The complete cone can only be obtained on emergence if all directions of vibration are represented in the incident ray

Fig 135 shows how the phenomenon of external conical refraction may be illustrated experimentally A plate of arragonite has its surfaces covered by opaque plates, each having an aperture If one of these plates be fixed and the other is movable, a position may be found of the apertures O and O' such that only such light can traverse the plate as passes along the axis of single ray velocity The rays on emergence are found to be spread out and to form the generating lines of a cone But as any ray after passing through a plate must necessarily be parallel to its original direction, it follows that to obtain the emergent cone, the incident beam must also be conical This may be secured by means of a lens LL' arranged as in the figure Those parts of the incident beam forming a solid cone which are not required, do not travel inside the crystal along OO' and hence are cut off by the plate covering the upper surface

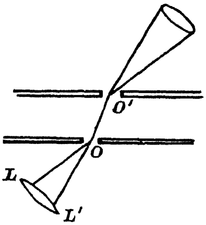

Fig 135

103 Fresnel's investigation of double refraction. Fresnel's method of treating double refraction which led him to the discovery of the laws of wave propagation in crystalline media, though not free from objection, is very instructive, and deserves consideration as presenting in a simple manner some of the essential features of a more complete investigation

Consider a particle P attracted to a centre O with a force a^2x when the particle lies along OX, and a force b^2y when it lies along OY The time of oscillation, if the particle has unit mass, is, by Art 2, $2\pi/a$ or $2\pi/b$ according as the oscillation takes place along the axis of X or along the axis of Y When the displacement has components both along OX and along

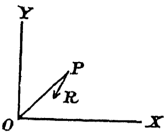

Fig 136

OY, the components of the force are a^2x and b^2y, and the resultant force is
$$R = \sqrt{a^4x^2 + b^4y^2}$$

The cosines of the angles which the resultant makes with the coordinate axes are a^2x/R and b^2y/R The direction of the resultant force is not the same as that of the displacement, the direction cosines of which are x/r and y/r The cosine of the angle included between the radius vector and the force is found in the usual way to be
$$\frac{a^2x^2 + b^2y^2}{Rr},$$
and the component of the force along the radius vector is
$$(a^2x^2 + b^2y^2)/r$$

If we draw an ellipse $a^2x^2 + b^2y^2 = k^2$, where k is a constant having the dimensions of a velocity, the normal to this ellipse at a point P, having coordinates x and y, forms angles with the axes, the cosines of which are in the ratio a^2x to b^2y, hence the force in the above problem acts in the direction ON of the line drawn from O at right angles to the tangent at P The component of the force along the radius vector is k^2/r, and the force per unit distance is k^2/r^2, so that if the particle were constrained to move on the radius vector OP, its period would be $2\pi r/k$ The ratio r/k depending only on the direction of OP our result is independent of the particular value we attach to k

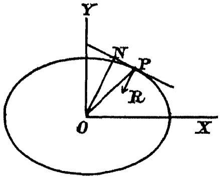

Fig 137

If we extend the investigation to three dimensions, the component of attraction along OZ being c^2z, we obtain the same result, and the component of force acting along any radius vector OP per unit length is k^2/r^2, where r is the radius drawn in the direction of OP to the ellipsoid
$$a^2x^2 + b^2y^2 + c^2z^2 = k^2.$$

If the displacement is in any diametral plane HPK of this ellipsoid (Fig 143), the normal PN does not in general lie in this plane, and the projection of PN on the plane does not pass through O, unless OP is a semiaxis of the ellipse HPK In the latter case, PL the tangent to the ellipse in the diametral plane, is at right angles to PO and to PN, and hence the plane containing PO and PN is normal to the plane of the section

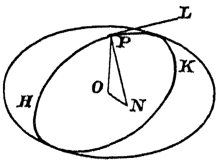

Fig 138

Fresnel considers the condition under which a plane wave propagation is possible in a crystalline medium The investigation in Art 11 has shown that the accelerations of any point in a plane

distortional wave of homogeneous type, are the same as those due to central attracting forces It is also clear that a plane polarized wave cannot be transmitted as a single wave unless the force of restitution is in the direction of the displacement If we disregard longitudinal waves as having no reference to the phenomena of light, we need only consider that component of the force which acts in the plane of the wave. This consideration leads to Fresnel's construction For if we take the ellipsoid

$$a^2x^2 + b^2y^2 + c^2z^2 = 1,$$

which, as we now see, is quite appropriately called the ellipsoid of elasticity, a central section parallel to the wave-front gives an ellipse which, by its principal axes, indicates the two directions of displacement which are compatible with a transmission of a single plane wave The periods of oscillation are proportional to the axes of this section, and as for a given wave-length the periods of oscillation are inversely proportional to the velocity of transmission, it follows that the velocities of the plane waves parallel to the section are inversely proportional to the axes of the ellipse of intersection We have thus arrived at the construction which has formed the starting point of our discussion of the phenomena of double refraction (Art 81)

The direction of the elastic force for any displacement being parallel to the normal to the ellipsoid of elasticity, drawn at the point at which the direction of the displacement intersects the ellipsoid, the proposition proved in Art 86 shows that the four vectors representing the direction of vibration, the elastic force, the ray and the wave-normal are coplanar

CHAPTER IX.

INTERFERENCE OF POLARIZED LIGHT

104 Preliminary Discussion If a plane unpolarized wave enters a plate of a doubly refracting substance, the two waves inside the crystal travel with different velocities and in slightly different directions, but on emergence both waves are refracted so as again to become parallel to their original directions If the wave was originally polarized in one of two definite directions, it is observed that there is only one refracted wave The two planes of polarization for which this is the case, are observed to be at right angles to each other. Inside the crystal the two waves are polarized in directions nearly though not quite at right angles to each other After emergence the planes of polarization are at right angles to each other, not only approximately, but strictly This follows from the principle of reversion, assuming the above-mentioned result of observation

In general, a wave of polarized light incident on a doubly refracting plate becomes polarized elliptically The axes of the ellipses vary with the wave-length and the thickness of material travelled through, hence also with the direction of the incident light, and the ellipse may, in particular cases, become a straight line or a circle If the emergent light is examined through a Nicol prism or any arrangement which transmits oscillations in one direction only, colour effects are observed which we shall have to explain in greater detail It is clear, however, that any interference effect must depend on the difference of phase in the two overlapping emergent waves Let *LO* (Fig 139) be an incident ray, forming part of a parallel beam, *OA* and *OB* the refracted *wave normals*, *AS* and *BT* the emergent wave normals Draw *AK* at right angles to *BT* and *BH* at right angles to *OA* Imagine a second incident ray, parallel to the first, and at such a distance that the wave normal which

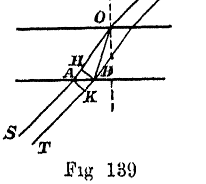

Fig 139

is parallel to OA passes through B, and is refracted outwards along BT then from the principle of wave transmission it follows that the optical length of BK is the same as that of AH In the emergent wave-front, the difference in optical length is therefore $\frac{OB}{v_1} - \frac{OH}{v_2}$, where v_1 and v_2 are the velocities of the waves along OB and OH respectively (The unit time is still taken to be such that the velocity of light in vacuo is one) The angle between OB and OA is small, and if we neglect its square, we may write $OB = OH$ The difference in optical length is therefore $\rho\left(\frac{1}{v_1} - \frac{1}{v_2}\right)$, where ρ is the length of that wave normal inside the plate, which lies nearest to the plate normal Unless the incidence is very oblique, it makes no difference, to the degree of approximation aimed at, along which wave normal ρ is measured, but for the sake of definiteness, we adhere to the specified meaning of ρ If OB and OA represent the refracted rays, we argue similarly that by Fermat's principle, optical lengths may be measured along a path near the real one, committing only an error of the second order The optical length for the ray of velocity s_2 might therefore be measured either along its real path OA or along its neighbour $OB + BK$, ending, of course, in the same wave-front We may therefore also express the difference in optical length as $t\left(\frac{1}{s_1} - \frac{1}{s_2}\right)$, where t is the length of ray inside the crystal and s_1, s_2, are the ray velocities. We may, according to convenience, use either one or the other two forms, which are both approximate only Which of these is the more accurate in a particular case depends on the question as to whether the angle between the two ray velocities or between the two wave normals is the smaller In the neighbourhood of the optic axes, it is preferable to refer the relative retardation to the wave normals

105. Intensity of illumination in transmitted light. Consider polarized light with its direction of vibration along OP (Fig 140), falling normally on the surface of a crystal which divides the wave into two portions, one vibrating along OX and one along OY After traversing the thickness of the plate, the two waves emerge normally with a difference of phase δ depending on the difference in optical length of the two wave normals inside the crystal If the amplitude of the incident light is one, the emergent waves have amplitudes $\cos\alpha$, $\sin\alpha$, if α is the angle between OP and OX, there being a difference in phase δ between them If now the emergent beam be examined through a Nicol prism called the "analyser," transmitting light only which vibrates along OA, the

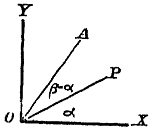

Fig 140

component k_1 of the transmitted light due to that portion which in the crystal had OX for its direction of vibration, is $k_1 = \cos\beta\cos\alpha$, similarly $k_2 = \sin\beta\sin\alpha$ is that component of the light which, having OY for the direction of vibration inside the crystal, is capable of traversing the analyser

Two rays of amplitude k_1 and k_2 and phase difference δ, polarized in the same direction, have a resultant, the intensity of which is

$$k_1^2 + k_2^2 + 2k_1k_2\cos\delta,$$

for which we may write

$$(k_1 + k_2)^2 - 4k_1k_2\sin^2\frac{\delta}{2}.$$

Substituting the values of k_1 and k_2 the intensity of the emergent beam becomes

$$I = \cos^2(\beta - \alpha) - \sin 2\alpha \sin 2\beta \sin^2\frac{\delta}{2} \tag{1}$$

All colour or interference effects shown by crystalline plates when examined by polarized light, depend on the application of this formula So long as there is only one parallel beam, the plate having the same thickness everywhere, all the quantities are constant, and the plate appears uniformly illuminated Important particular cases are those in which the Nicols are either parallel ($\alpha = \beta$), or crossed at right angles $\left(\beta - \alpha = \pm\frac{\pi}{2}\right)$

In the first case we have

$$I_0 = \left(1 - \sin^2 2\alpha \sin^2\frac{\delta}{2}\right),$$

and in the second $$I_1 = \sin^2 2\alpha \sin^2\frac{\delta}{2},$$

which shows that $$I_0 + I_1 = 1.$$

This relation is a particular case of the general law that if for any value of α and β, $I = I_A$, and if I becomes I_B on turning either the analyser or polarizer through a right angle, then $I_A + I_B = 1$

We may convince ourselves that this is true without having recourse to the equations The light falling on the analysing Nicol is partly transmitted and partly deviated to one side, the two portions making up together the incident light which is supposed to be white On rotating the Nicol through a right angle the transmitted and deviated portions are interchanged so that the complementary effect must be observed

When white light passes through the plate, the relative proportion of different colours is not in general preserved because δ depends on the wave-length If a is the amplitude of light of a particular wave-length, so that white light may be represented by Σa^2, the light transmitted through the system is

$$\cos^2(\alpha-\beta)\,\Sigma a^2 - \sin 2\alpha \sin 2\beta\, \Sigma\left(a^2 \sin^2\frac{\delta}{2}\right)$$

The first term represents white light of intensity proportional to $\cos^2(\alpha-\beta)$, and the second term represents coloured light The relative proportion of the different wave-lengths is not affected by a change in α or β, but the total colour effect may change because the product $\sin 2\alpha \sin 2\beta$ may be either positive or negative In the first case, we get a certain colour, in the second, white light minus that colour, *i e* the complementary colour We distinguish two special cases.

Case 1 The Nicols are crossed so that $\alpha-\beta=\frac{\pi}{2}$ Here we have

$$I = \sin^2 2\alpha\, \Sigma\left(a^2 \sin^2\frac{\delta}{2}\right)$$

The colours are most saturated in this case, because there is no admixture of white light As the axes of x and y are fixed in the crystal, we may vary α without change of $\alpha-\beta$ by turning the crystalline plate in its own plane There will then be four places of maximum intensity at which $\alpha=45^\circ$ or an odd multiple thereof, and four places of zero intensity at which α is a multiple of 90°

Case 2 The Nicols are parallel so that $\alpha=\beta$ Here we have

$$I = \Sigma a^2 - \sin^2 2\alpha\, \Sigma\left(a^2 \sin^2\frac{\delta}{2}\right)$$

The colour here is always complementary to that in the previous case for the same value of α, the light being white when α is a multiple of a right angle, and most saturated when α is an odd multiple of 45°.

If for any value of α and β, the crystal is turned in its own plane, there are eight positions at which $\sin 2\alpha \sin 2\beta$ vanishes, these occur whenever one of the axes OX and OY coincides with the principal planes of either the polarizing or analysing Nicol In these positions of the crystal, the light is white, and on passing through these positions, the colour changes into its complementary

106. Observations of colour effects with parallel light. The general experimental arrangement by means of which the colour effects of polarized parallel light may be shown, is sketched dia

grammatically in Fig 141 MM' is a mirror reflecting the light from the sky, N_1 and N_2 the polarizing and analysing Nicols, CC' is the crystalline plate

The field of view of a Nicol prism is much restricted by the increased distance of the eye from the polarizer N_1 Hence when light from a distant source, such as the sky, passes through both Nicols, only such waves reach the eye as subtend a small angle The eye at E, focussed for infinity, receives light therefore which has passed through the crystal nearly in the normal direction, and the crystal appears coloured with a uniform tint If the eye is focussed on the crystal, the colours are not so pure because the different rays leaving the same point of the crystal have traversed it at different inclinations, but when the crystal is thin, so that the relative retardation is only a few wave-lengths, a small variation in direction does not produce much effect on the colour, and therefore the colours are seen with the eye focussed on the plate, nearly as well as with the eye adjusted for parallel light

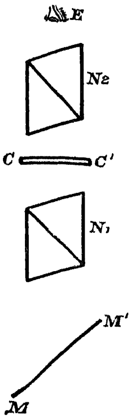

Fig 141

An interesting variation of the experiment may be made if the analysing Nicol is replaced by a double image prism, two partially overlapping images of the plate are then seen The images are coloured where they are separate, but white where they overlap, showing that the colours are complementary

107 Observations with light incident at different angles If the field of view is enlarged so as to include rays which have traversed the crystal at sensibly different angles, the effects are more complicated because they depend on the part of the crystal looked at, so that the plate appears to be covered with a pattern of coloured bands To realize experimentally the necessary increase of the field of view, we may look at the crystal plate through an inverted telescopic system consisting of two lenses L_1 and L_2, placed so as to *diminish* angular distances The different parallel pencils which have passed through the crystal, pass out of this system with their axes more nearly parallel, so that they may now be sent through a Nicol A similar telescopic system $K_2 K_1$ serves to *increase* the angular deviation of the rays which

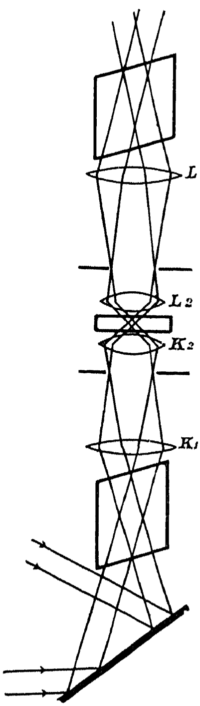

Fig 142

have passed through the polarizing Nicol The thickness of the plate used ought now to be rather larger because it is desired to bring out the differences which are due to variations of length of paths and inclination When crystals are examined in this fashion, it is generally said that convergent or divergent light is used, but it must be clearly understood that the rays of light which are brought together on the retina traversed the crystal as a parallel pencil So long as the eye is focussed for infinity, the sole distinction between this case and the previous one, lies in the increase of the field of view

108 Uniaxal Plate cut perpendicularly to the axis. In order to show how the equation (1) is to be applied to the explanation of the interference pattern under the experimental conditions of the last article, we may treat first the simple case of a plate cut normally to the axis of a uniaxal crystal An eye E looking in an oblique direction through such a plate (Fig 143) receives rays which have passed through lengths of path in the crystal, which only depend on the angle between the line of vision and the normal to the plate Hence the retardation δ is the same along a circle drawn on the surface of the crystal, having its centre coincident with the foot of the perpendicular from the eye to the plate As the colour effects depend on δ, the field of view is traversed by coloured circular rings A line along which δ is constant is called an isochromatic line, but the term *isochromatic* here includes the complementary colour The illumination is not constant along an isochromatic line on account of the variations of α and β In Fig 144 $ABCD$ represents the plate, N the foot of the perpendicular from the eye to the plate If the line of vision passes through the point O, NO is the trace of the plane of incidence, and this plane also contains the optic axis. The two directions of vibration of the ray inside the crystal are therefore NO and the line at right angles to it, and to make equation (1) apply, we must put the axes of X and Y along those directions The circle drawn through O with N as centre is an isochromatic line The polarizing and analysing directions remain fixed in space, while the coordinate axes revolve with the point O round N. Whenever either $\sin 2\alpha = 0$ or $\sin 2\beta = 0$, the colour term disappears and we obtain therefore in general four diameters along which there is no coloration The lines drawn along these directions are called achromatic lines

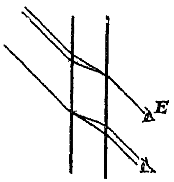

Fig. 143

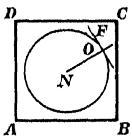

Fig. 144

We consider three cases

Case 1. The Nicols are crossed, *i e* $\alpha - \beta = \frac{\pi}{2}$

The intensity as before is given by

$$I = \sin^2 2\alpha \, \Sigma \left(a^2 \sin^2 \frac{\delta}{2} \right).$$

There are two lines at right angles to each other, along which the intensity is zero, these lines coinciding with the directions of the planes of polarization of the analysing and polarizing Nicols The intensity is greatest at an angle of 45° from these lines The field is traversed by rings of varying colours, or in the case of homogeneous light, by coloured rings of varying intensity, the dark rings corresponding to the positions at which the phase retardation δ is a multiple of four right angles The whole appearance consists therefore of a number of concentric rings with a dark cross, as shown in the photograph reproduced in Plate II, Fig 1 The cross widens out away from the centre and each of its branches is sometimes referred to as a "brush."

Case 2 The Nicols are parallel, *i.e* $\alpha = \beta$

The intensity is

$$\Sigma a^2 - \sin^2 2\alpha \, \Sigma \left(a^2 \sin^2 \frac{\delta}{2} \right),$$

and the whole effect is complementary to that observed in the first case The rings are now crossed by *bright* brushes Plate II, Fig 2 shows the appearance

Case 3 This includes all positions of the analyser and polarizer in which these are neither parallel nor crossed There are four achromatic lines corresponding to $\alpha = 0$ and $\alpha = \frac{\pi}{2}$, $\beta = 0$ and $\beta = \frac{\pi}{2}$

Along an isochromatic circle, the colour changes into its complementary (or for homogeneous light, a minimum of light changes into a maximum) on crossing one of the achromatic lines This is shown in Plate II, Fig 3 which is also a reproduction of a photograph When either the axis of x or the axis of y falls within the acute angle formed by the directions of the analyser and polarizer, the product $\sin 2\alpha \sin 2\beta$ is negative so that the maxima of light are brighter and the minima less dark The field is therefore separated into segments of unequal illumination and may at first sight give the fictitious appearance of a dark cross. The eight achromatic brushes in this case separate the bright and dark segments, and are not very conspicuous

109 Relation between wave velocities In order to discuss the form of the achromatic and isochromatic lines in more complicated cases, it is necessary to calculate the phase difference δ The first step consists in finding an expression for $\frac{1}{v_1} - \frac{1}{v_2}$, the difference between the

reciprocals of the wave velocities of plane waves travelling in the same direction It was proved in Art 90 that the axes of the elliptical section of the ellipsoid of elasticity are the bisectors of OQ and OQ' Fig 114, Q and Q' being in the planes containing the wave normal ON and optic axes OH, OH'

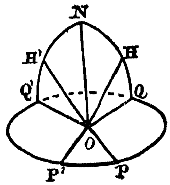

Fig 114

The radii OP and OP' at right angles to OQ, OQ' and therefore also at right angles to the optic axes, belong to the circular sections, and have therefore a length $1/b$ If v_1 and v_2 are the reciprocals of the semiaxes of the ellipse, and the angle POP' is denoted by 2ϕ, we have from the equation to the ellipse, reduced to polar coordinates,

$$v_1^2 \cos^2 \phi + v_2^2 \sin^2 \phi = b^2$$

The angle A between OQ and OQ', and the angle 2ϕ, are supplementary to each other, so that

$$\cos^2 \phi = \sin^2 \frac{A}{2} = \frac{1}{2}(1 - \cos A),$$

$$\sin^2 \phi = \cos^2 \frac{A}{2} = \frac{1}{2}(1 + \cos A),$$

$$(v_1^2 + v_2^2) + (v_2^2 - v_1^2) \cos A = 2b^2$$

Another relation between the wave velocities is obtained by making use of the fact that in any ellipsoid the sum of the reciprocals of the squares of any three diameters at right angles to each other, is constant. The section we are considering has $1/v_1$ and $1/v_2$ for semiaxes, and l, m, n for the direction cosines of its normal Hence $l^2a^2 + m^2b^2 + n^2c^2$ is the reciprocal of the square of the radius vector, which is normal to the section, and

$$v_1^2 + v_2^2 + l^2a^2 + m^2b^2 + n^2c^2 = a^2 + b^2 + c^2 \qquad (2),$$

or by making use of $\quad l^2 + m^2 + n^2 = 1,$

$$(v_1^2 + v_2^2) = (a^2 + c^2) + l^2(b^2 - a^2) + n^2(b^2 - c^2),$$

$$(v_1^2 - v_2^2) \cos A = (a^2 + c^2 - 2b^2) + l^2(b^2 - a^2) + n^2(b^2 - c^2)$$

$$= (a^2 - b^2)(m^2 + n^2) + (c^2 - b^2)(l^2 + m^2) \qquad (3)$$

The expressions are simplified still further if, instead of A, we introduce the angles θ_1 and θ_2 between ON and OH, OH' respectively, and the angle σ included between the two optic axes

The spherical triangle formed by OH, OH', ON gives

$$\cos A = \frac{\cos \sigma - \cos \theta_1 \cos \theta_2}{\sin \theta_1 \sin \theta_2}$$

The optic axes lie in the plane xz and if their positive directions are

chosen so that the axis of X bisects the angle included between them, we find, writing l_1, n_1, l_2, n_2 for their direction cosines

$$\cos\theta_1 = ll_1 + nn_1,$$

$$\cos\theta_2 = ll_1 - nn_1,$$

$$\cos\theta_1\cos\theta_2 = l^2l_1^2 - n^2n_1^2$$

For the angle σ between the optic axes we have

$$\cos\sigma = l_1l_2 + n_1n_2 = l_1^2 - n_1^2,$$

$$\sin\theta_1\sin\theta_2\cos A = \cos\sigma - \cos\theta_1\cos\theta_2$$

$$= l_1^2(m^2+n^2) - n_1^2(l^2+m^2),$$

and introducing the values of l_1 and n_1, Art 85,

$$= \frac{(a^2-b^2)(m^2+n^2) - (b^2-c^2)(l^2+m^2)}{a^2-c^2}$$

Comparing this with (3), we finally obtain the simple equation

$$v_1^2 - v_2^2 = (a^2-c^2)\sin\theta_1\sin\theta_2 \qquad (4)$$

We may also transform

$$\cos\theta_1\cos\theta_2 = l^2l_1^2 - n^2n_1^2 = \frac{l^2(a^2-b^2) - n^2(b^2-c^2)}{a^2-c^2},$$

$$v_1^2 + v_2^2 = (a^2+c^2) - (a^2-c^2)\cos\theta_1\cos\theta_2 \qquad (5)$$

Combining (4) and (5) we may express separately

$$2v_1^2 = (a^2+c^2) - (a^2-c^2)\cos(\theta_1+\theta_2)$$

$$2v_2^2 = (a^2+c^2) - (a^2-c^2)\cos(\theta_1-\theta_2).$$

If the difference between a and c is so small that its square may be neglected, we may write

$$\frac{1}{v_2} - \frac{1}{v_1} = \frac{v_1 - v_2}{v_1v_2} = \frac{v_1^2 - v_2^2}{2v^3},$$

where v stands for the velocity to which both v_1 and v_2 approach when $a - c$ vanishes For v we may therefore write either $\sqrt{ac}$ or $\frac{1}{2}(a+c)$, and for $2v^3$ we may write $ac(a+c)$

Introducing the values of $v_1^2 - v_2^2$ from (4) we obtain

$$\frac{1}{v_2} - \frac{1}{v_1} = \frac{(a^2-c^2)}{(a+c)ac}\sin\theta_1\sin\theta_2,$$

or

$$\frac{1}{v_2} - \frac{1}{v_1} = \left(\frac{1}{c} - \frac{1}{a}\right)\sin\theta_1\sin\theta_2 \qquad (6)$$

110. Relation between ray velocities. The proposition contained in the last article represents a theorem which may be applied to any ellipsoid of semiaxes $1/a$, $1/b$, $1/c$, if v_1, v_2 are the reciprocals of the principal axes of a section which forms angles θ_1 and θ_2 with the circular sections We may therefore write down at

once the corresponding equations for the reciprocal ellipsoid, substituting the ray velocities s_1 and s_2 for $1/v_1$ and $1/v_2$. We obtain in this way

$$2s_1^{-2} = (a^{-2} + c^{-2}) - (a^{-2} - c^{-2}) \cos(\eta_1 + \eta_2),$$
$$2s_2^{-2} = (a^{-2} + c^{-2}) - (a^{-2} - c^{-2}) \cos(\eta_1 - \eta_2),$$
$$s_1^{-2} + s_2^{-2} = (a^{-2} + c^{-2}) - (a^{-2} - c^{-2}) \cos\eta_1 \cos\eta_2,$$
$$s_1^{-2} - s_2^{-2} = (a^{-2} - c^{-2}) \sin\eta_1 \sin\eta_2,$$

where η_1 and η_2 are the angles formed between the normal to the section and the axes of single ray velocities.

111. The surface of equal phase difference, or Isochromatic Surface. If we imagine a number of plane waves crossing at a point O (Fig 145) in a crystalline medium, there being two wave velocities in each direction, we may construct a surface such that at any point P, belonging to the surface, the phase difference δ between the two wave-fronts which have OP for wave normals is the same. If ρ be any radius vector OP, v_1 and v_2 the wave velocities, the two optical distances from O to P are ρ/v_2 and ρ/v_1, hence the required surface has for equation

$$\rho\left(\frac{1}{v_2} - \frac{1}{v_1}\right) = \text{constant}$$

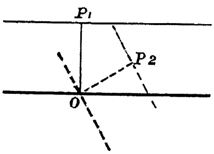

Fig 145

It will be sufficient to confine the discussion to the case of a small difference between the two wave velocities. We shall consider therefore $a - c$ and a fortiori $a - b$ to be so small that their squares may be neglected. We may then apply equation (6) and by introducing the principal indices of refraction $\mu_1 = \frac{1}{a}$, $\mu_3 = \frac{1}{c}$, the equation to the surface of equal phase difference is obtained in the form

$$\rho(\mu_3 - \mu_1) \sin\theta_1 \sin\theta_2 = \delta \qquad (7).$$

Unless highly homogeneous light is used, δ must not exceed a small multiple of a wave-length, if interference effects are to be observed. It follows that unless the observations are carried out close to one of the optic axes, in which case either $\sin\theta_1$ or $\sin\theta_2$ is small, $\mu_3 - \mu_1$ must be small. This justifies the simplification we have introduced in treating $a - c$ as a small quantity.

In uniaxal crystals, there is only one axis, so that putting $\theta_1 = \theta_2 = \theta$, the polar equation to the surface of equal phase difference or "isochromatic" surface then becomes

$$\rho(\mu_e - \mu_0) \sin^2\theta = \delta \qquad (8)$$

This surface is formed by the revolution about the optic axis of a family of curves for which the polar equation is represented by

(8) and which is drawn to scale in Fig 146 Only half of the curves is shown, there being symmetrical halves below the line PQ The scale is such that if the substance is Iceland Spar, and the length

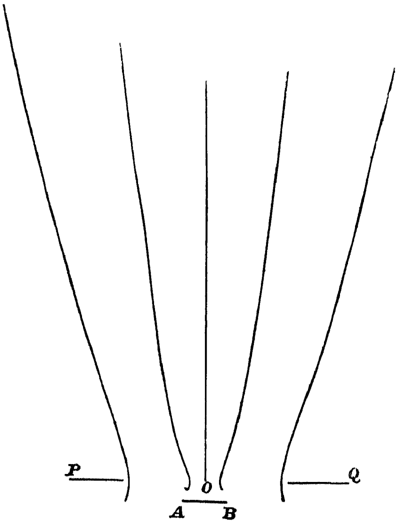

Fig 146

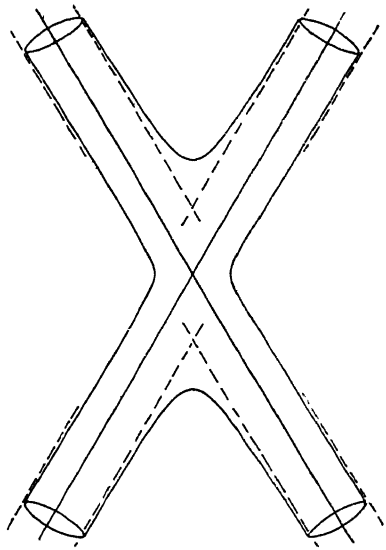

Fig 147

marked AB represents one millimetre, the inner curve is the isochromatic surface of phase difference equal to 100 wave-lengths, the wave-length being that of sodium light, the phase difference belonging to the outer curve is five times as great OC is the optic axis The upper portions of the curves are sensibly parabolic, because when θ is small, the radius vector ρ is nearly equal to its projection x on the optic axis, so that the equation to the curve becomes

$$z^2/x = \text{constant}$$

In biaxal crystals the isochromatic surface has four sheets surrounding the optic axes. Their intersection with the plane containing these axes is represented in Fig 147 for the case where the angle between the optic axes is 60° When ρ is infinitely large, it follows from (7) that either θ_1 or θ_2 is zero. If θ_1 vanishes, θ_2 must be equal to σ, the angle included between the optic axes. For large values of ρ we may still take approximately $\theta_2 = \sigma$ and the equation to the isochromatic surface approaches therefore a surface the equation to which is by (7)

$$\rho \sin \theta_1 = \delta \operatorname{cosec} \sigma/(\mu_2 - \mu_1)$$

This is the equation to a circular cylinder, having one of the optic axes as axis. The intersection of this cylinder with the plane of the paper

gives two straight parallel lines, which are the asymptotes to the curve which forms the intersection of the isochromatic surface with the plane containing the optic axes If ρ' be the distance of the asymptotes from the origin

$$\rho' = \delta \operatorname{cosec} \sigma/(\mu_3 - \mu_1),$$

there are two similar asymptotes parallel to the second optic axis These asymptotes are shown by dotted lines in the figure, and it will be noticed that each of them intersects one branch of the curve to which it is a tangent at infinity

The two distances ρ_0 and ρ_1 of the vertices of the surface may be found by substituting $\theta_1 = \theta_2 = \frac{1}{2}\sigma$ and $\theta_1 = \frac{\pi - \sigma}{2}$, $\theta_2 = \frac{\pi + \sigma}{2}$ respectively We then find

$$\rho_0 = \delta \operatorname{cosec}^2 \tfrac{1}{2}\sigma/(\mu_3 - \mu_1),$$
$$\rho_1 = \delta \sec^2 \tfrac{1}{2}\sigma/(\mu_3 - \mu_1)$$

112. Application of the Isochromatic Surface to the study of polarization Let a doubly refracting plate, Fig 148, receive light at different inclinations An eye placed at E and looking towards a point S on the plate observes certain interference effects Tracing the disturbance backwards from E, there will be two wave normals within the plate corresponding to SE Let $OS = \rho$ be that wave normal which forms the smaller angle with OM the normal to the plate According to Art 104 the difference in path at S, and therefore at E, of the two waves which have traversed the crystal is $\rho\left(\frac{1}{v_1} - \frac{1}{v_2}\right)$ A similar reasoning applies to the interference observed in the direction ET, O_1T being the direction of the wave normal inside the crystal Draw OS_1 parallel to O_1T and EN at right angles to the plate The interference seen at T is the same as that due to the phase difference at S_1 for waves propagated through O

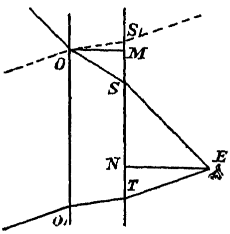

Fig 148

If $\imath$ be the inclination of the line of sight

$$\tan \imath = \frac{NS}{NE}$$

If r is the inclination of OS to OM

$$\tan r = \frac{MS}{MO},$$

$$\frac{NS}{MS} = \frac{NE}{MO} \frac{\tan \imath}{\tan r}$$

For small values of i the ratio $\tan i/\tan r$ is nearly constant even in the case of doubly refracting crystals (for which $\sin i/\sin r$ is not, strictly speaking, constant). Representing this ratio by μ

$$\frac{NS}{MS} = \mu \frac{NE}{MO},$$

and similarly

$$\frac{NT}{MS_1} = \mu \frac{NE}{MO}$$

If an isochromatic surface be constructed with O as centre, it follows that its intersection with the upper surface of the plate enlarged in the ratio $\mu NE/MO$ gives the interference pattern as it is seen projected on the plate. When $\tan i/\tan r$ is not constant, there is a certain distortion due to the variability of that factor.

As an example we may use Fig. 146 to construct the isochromatic lines for a plate of Iceland spar. Place the plate with its normal in the plane of the paper, its lower surface passing through O with OC along the optic axis. The upper surface will intersect the plane of the paper in a line which is at a distance from O equal to the thickness of the plate, the length of AB representing one millimetre. The intersection of the isochromatic surfaces which are formed by the revolution of the curves drawn in the figure about OC and the upper surface of the plate, will show the isochromatic lines for a phase difference of 100 and 500 wave-lengths. As all isochromatic surfaces may be obtained from one by increasing the length of the radius vector in a given proportion, we may obtain all isochromatic curves from the same surface by changing the scale. Thus to obtain the curve for which the retardation is ten wave-lengths, in the above example, we must, taking the inner curve, alter the scale, so that AB represents 1 mm. In simple cases, this method of forming a rapid idea of the shape of the interference curves is very serviceable, the different curves being obtained by drawing the upper surface of the plate at different distances from the origin.

113. Isochromatic curves in uniaxal crystals. To study the intersection of the isochromatic surface by a plane drawn in any direction, construct a spherical triangle SNO (Fig. 149), such that if C be the centre of the sphere, CN is parallel to the normal of the plate, CO parallel to the optic axis, and CS parallel to any wave normal inside the plate. Also let

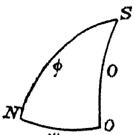

Fig. 149.

θ = angle between CS and CO,

ϕ = „ „ CN and CS,

ψ = „ „ CN and CO,

A = „ „ planes CNS and CNO.

In the spherical triangle NOS

$$\cos\theta = \cos\psi\cos\phi + \sin\psi\sin\phi\cos A\,,$$

$$\sin^2\theta = \sin^2\psi + \sin^2\phi\,(\cos^2\psi - \sin^2\psi\cos^2 A) - \sin 2\psi\sin\phi\cos\phi\cos A$$

To obtain the isochromatic curves we must take the intersection between the surfaces given by (8)

$$\rho\sin^2\theta = \text{constant}$$

and the plane at which $\rho\cos\phi = e,$

where e is the thickness of the plate

Eliminating ρ we obtain an equation for the curves in the form

$$\frac{\sin^2\theta}{\cos\phi} = \text{constant} \tag{9}$$

We shall consider the angle of internal incidence to be so small that we may write sensibly $\dfrac{1}{\cos\phi} = 1 + \frac{1}{2}\sin^2\phi$, and rejecting all terms involving a higher power of ϕ than the second

$$\frac{\sin^2\theta}{\cos\phi} = \sin^2\psi + \sin^2\phi\,(\tfrac{1}{2} + \tfrac{1}{2}\cos^2\psi - \sin^2\psi\cos^2 A) - \sin 2\psi\sin\phi\cos A$$

An important special case occurs when the plate is cut parallel to the axis In that case $\sin\psi = 1$ and $\sin 2\psi = 0$ so that the condition for equality of phase difference at the upper surface becomes

$$\sin^2\phi\,(\tfrac{1}{2} + \tfrac{1}{2}\cos^2\psi - \sin^2\psi\cos^2 A) = \text{constant},$$

or introducing the value of ψ

$$\sin^2\phi\,(\sin^2 A - \cos^2 A) = \text{constant}$$

If we introduce rectangular coordinates with the pole of the plate normal N as centre, so that

$$x = e\sin\phi\sin A$$

$$y = e\sin\phi\cos A,$$

the equation to isochromatic curves reduces to

$$x^2 - y^2 = \text{constant}$$

These curves are therefore rectangular hyperbolas, one of the axes being parallel to the direction of the optic axis and the other at right angles to it If $\sin 2\psi$ does not vanish, then for small values of ϕ the term involving the first power of ϕ is the important one Close to the normal therefore in a plate cut obliquely to the axis, the isochromatic lines are given by

$$\sin\phi\cos A = \text{constant},$$

which represents straight lines at right angles to the plane containing the optic axis and normal When ϕ becomes sufficiently large for the

second order terms to become appreciable, these lines become curved, but both terms together still represent conic sections.

Unless the normal to the plate is nearly coincident with the optic axis, there are no achromatic lines, as the axes of x and y remain sensibly parallel throughout the field

114. Isochromatic Curves in Biaxal Crystals We shall not follow out in detail the calculation in this case, but only indicate the method which may conveniently be adopted Construct the spherical triangle NSO_1, NSO_2, Fig 150, corresponding to NSO, Fig 149, only with the difference that we have now two optic axes CO_1 and CO_2 Let A represent the angle between the sides NS and NG, where NG is a large circle bisecting the angle between NO_1 and NO_2. If ω be half the angle between the sides NO_2 and NO_1, we have in the triangle NO_2S

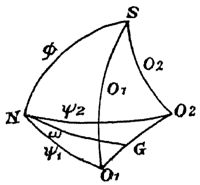

Fig 150.

$$\cos\theta_2 = \cos\psi_2\cos\phi + \sin\psi_2\sin\phi\cos(A-\omega),$$

and similarly in the triangle NO_1S

$$\cos\theta_1 = \cos\psi_1\cos\phi + \sin\psi_1\sin\phi\cos(A+\omega)$$

The angle 2ω may be obtained from ψ_1, ψ_2, and O_1O_2, the angle between the optic axes.

From the above two equations we may obtain

$$\sin^2\theta_1\sin^2\theta_2/\cos^2\phi,$$

expressed in a series proceeding by ascending powers of $\sin\phi$

It is found that when the normal of the plate coincides with one of the axes of elasticity, the factor of the first power of $\sin\phi$ is zero, and in that case, neglecting $\sin^4\phi$, the condition for the isochromatic lines is obtained by putting the factor of $\sin^2\phi$ equal to zero We thus obtain, as in the last article, the equation of rectangular hyperbolas. When the plate is cut obliquely the factor of $\sin\phi$ is the important one, and the curves close to the normal are straight lines, as with uniaxal crystals*

115. Biaxal Crystals cut at right angles to the bisector of the angle between the optic axes. This case has special interest, and may be treated in a very simple manner, if the angle between the optic axes is small Let OM_1, OM_2 be the directions of

* For the details of working out the general case, see Kirchhoff, *Vorlesungen uber mathematische Optik*, p 256

the optic axes. When these nearly coincide with the normal, the angles OM_1P and OM_2P are nearly right angles, so that approximately,

$$\sin\theta_1 = \frac{PM_1}{OP}, \quad \sin\theta_2 = \frac{PM_2}{OP}$$

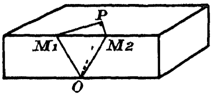

Fig 151.

Hence the equation to the isochromatic curve is

$$\frac{PM_1 \times PM_2}{OP} = \text{constant}$$

If further, OP form a small angle with the normal, we may consider it to be constant and equal to the thickness of the plate The isochromatic lines are in that case the lines on the surface of the plate which satisfy the equation

$$r_1 r_2 = \text{constant},$$

where r_1 and r_2 are measured from the points M_1 and M_2 on the crystal, such that plane waves traced back along the lines of vision EM_1 and EM_2 (Fig 152) are refracted with their wave-normals parallel to the optic axes The curves are so-called lemniscates For small values of the constants they split up into separate curves, each surrounding one of the points M_1 or M_2. For large values of the constants, they are nearly circular, with the point halfway between M_1 and M_2 as centre Figs 4 and 5, Plate II., shew the appearance.

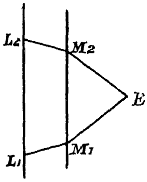

Fig 152

116 The Achromatic Lines in Biaxal Crystals. To trace the achromatic lines in a biaxal crystal cut so that the surface of the plate forms equal angles with the optic axes, we must introduce the condition that $\sin 2\alpha$ or $\sin 2\beta$ is zero The planes of polarization of the wave having OP (Fig 151) for normal bisect internally and externally the angles between the planes POM_1 and POM_2 (Art 90), and if OM_1, OM_2 and OP are all at a small inclination to the normal of the plate, the planes of polarization intersect the upper surface of the crystalline plate in lines which are very nearly the internal and external bisectors of the angle between the lines PM_1 and PM_2. Let the plane of the paper (Fig 153) represent the upper surface of the crystal, N being the foot of the perpendicular from O to that surface. Let also NY represent the principal plane of the polarizer or of the analyser The condition for the achromatic line implies that the bisector PT of M_1PM_2 must be parallel or at right angles to the fixed direction NY Let the coordinates of P be x and y; those of M_2,

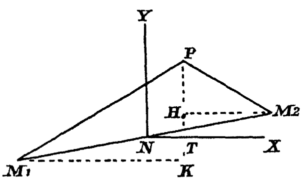

Fig 153

PLATE II

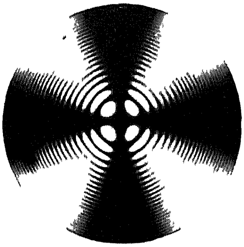

FIG. 1

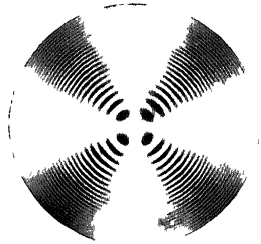

FIG. 2

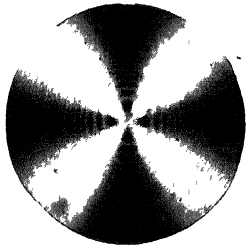

FIG. 3

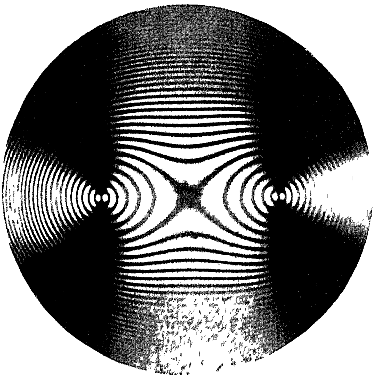

FIG. 4

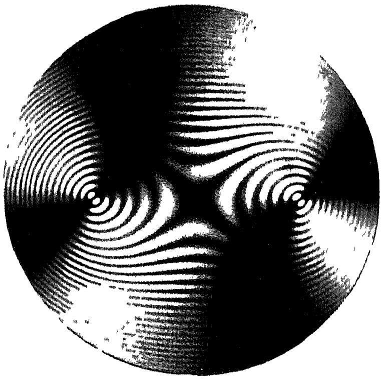

FIG. 5

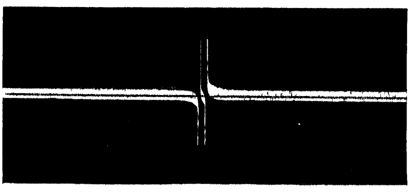

FIG. 6

[To face page 194

ξ and η, those of M_1, $-\xi$ and $-\eta$ Draw M_2H and M_1K at right angles to the bisector PT Then

$$\tan HPM_2 = \frac{HM_2}{HP}$$

and

$$\tan KPM_1 = \frac{KM_1}{KP}$$

Hence

$$\frac{KP}{HP} = \frac{KM_1}{HM_2},$$

$$\frac{y+\eta}{y-\eta} = \frac{\xi + x}{\xi - x},$$

$$\therefore \frac{y}{\eta} = \frac{\xi}{x},$$

$$xy = \xi\eta$$

Hence the achromatic lines are equilateral hyperbolas passing through the points M_1 and M_2 Fig 4, Plate II, shows these curves in the form of dark hyperbolic brushes.

If $\xi = \eta$, OY is at an angle of 45° to the line joining M_1 and M_2, and these two points are on the vertices of the curves This case is represented in Plate II, Fig 4

If either $\xi = 0$ or $\eta = 0$, $xy = 0$, and the achromatic lines coincide with the axes of x and y As in the case of uniaxal crystals, there are two sets of achromatic lines, one belonging to the polarizer and one to the analyser Both have the same shape, and both pass through the same points M_1 and M_2 If the observing Nicols are crossed or parallel, the two sets of lines coincide, in the former case the hyperbolic brushes are dark, and in the latter, bright Plate II, Fig. 5, gives a second example of dark hyperbolic brushes

117. Measurement of angle between optic axes. The intersection of the isochromatic surface (Fig 147) with planes drawn at different distances from O, shows that for small differences of path the interference rings surround the optic axes in closed curves This affords a means of determining the angle between the optic axes If a plate of a crystal cut symmetrically to the axes, as assumed in the last two articles, be mounted so that it can be rotated about an axis at right angles to the axes through an angle which can be measured, we may bring first one centre of the ring system belonging to one optic axis against a fixed mark in the observing telescope, and then the centre of the system belonging to the other axis The angle of rotation is the so-called "apparent angle" between the optic axes, for it is clear that what is measured is the angle between the lines of vision M_1E and M_2E (Fig 152). This angle is to be corrected for refraction to get the angles formed between L_1M_1 and L_2M_2

118. Dispersion of Optic Axes. We have treated the problems of double refraction as if the position of the optic axes were independent of the wave-length. Though the position of the principal axes does not in most cases depend on the wave-length, the principal velocities are different for the different colours. Now e_1, e_2, e_3, being the principal velocities for one wave-length, and e_1', e_2', e_3', for another, the latter quantities are not in general proportional to the former, hence the positions of the optic axes change with the colour of the light In some crystals the difference is very considerable

119. Two plates of a uniaxal crystal crossed. A great variety of effects may be produced by allowing light to traverse several plates in succession We shall only consider one case, which is of some importance

Let a plate be cut obliquely to the axis of a uniaxal crystal, and then divided into two halves which are therefore necessarily of the same thickness Superpose the two halves and turn one of them through a right angle We shall determine the shape of the isochromatic lines in this case

The first plate produces a difference in optical length between two coincident wave normals, which as obtained from (9) is

$$\delta_1 = (\mu_0 - \mu_e)\frac{e \sin^2 \theta_1}{\cos \phi},$$

the meaning of the letters being the same as that of Article (113). The second plate being turned through a right angle, the direction of vibration in the ordinary and extraordinary rays is interchanged, so that the phase difference in that plate is

$$\delta_2 = (\mu_e - \mu_0)\frac{e \sin^2 \theta_2}{\cos \phi'}.$$

The values of $\cos \phi$ and $\cos \phi'$ are nearly equal for the double reason that ϕ is small, and that the difference between μ_e and μ_0 is small Hence the total phase difference is proportional to

$$\sin^2 \theta_2 - \sin^2 \theta_1 = \cos^2 \theta_1 - \cos^2 \theta_2.$$

According to Art 113

$$\cos \theta_1 = \cos \phi \cos \psi + \sin \phi \sin \psi \cos A.$$

To find the angle θ_2 which the optic axis makes with the plate normal in the upper plate, we have only to increase the angle A by a right angle, keeping all other quantities the same. Hence

$$\cos \theta_2 = \cos \phi \cos \psi - \sin \phi \sin \psi \sin A.$$

Neglecting higher powers of $\sin \phi$

$$\cos^2 \theta_1 - \cos^2 \theta_2 = \sin \phi \sin 2\psi (\cos A + \sin A).$$

Introducing rectangular coordinates, so that

$$e \sin\phi \cos A = x, \quad e \sin\phi \sin A = y,$$

the equation to the isochromatic line for which the total difference in optical length $\delta_1 + \delta_2$ is equal to $n\lambda$, becomes

$$(\mu_e - \mu_0) \sin 2\psi\, (x + y) = n\lambda \tag{10}$$

This represents a series of parallel lines The field of view is therefore crossed by a series of bands, the central one not being coloured The bands are the wider apart the smaller ψ, so that if the bands are to be broad, the plate should be cut nearly normally to the optic axis It is found that in this case, the departure from straightness which depends on terms involving $\sin^2\phi$ is also small

Two plates combined together in the manner described, form the essential portions of the "Savart" polariscope, which is the most delicate means we possess for detecting polarized light The double plate is provided with an analyser, consisting of a Nicol prism or a Tourmaline plate In both cases, the plane of transmittance through the analyser should bisect the angle between the principal planes of the Savart plates in order to get the most sensitive conditions If the incident light be polarized at right angles to the plane of transmittance, the eye sees a dark central band accompanied on both sides by parallel coloured fringes If the incident light be polarized parallel to the direction which can pass through the analyser, the central band is bright, and the whole effect is complementary to that observed in the previous case By examining the light reflected from the sky or from almost any surface, the coloured fringes are noticed, and by rotating the whole apparatus we may find the direction in which the fringes are most brilliant and hence determine the plane of polarization of the incident light

120. The Half Wave-length Plate. If plane polarized light falls normally on a plate of a crystal cut to such a thickness that the two waves are retarded relatively to each other by half a wave-length, or a multiple thereof, the transmitted beam is plane polarized Let OX and OY be the two principal directions of vibration in the crystal, and α the angle between OX and the direction of vibration of the incident beam The displacements resolved along OX and OY may then be expressed by

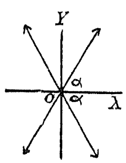

Fig 154

$$u = a \cos\alpha \cos\omega t,$$
$$v = a \sin\alpha \cos\omega t$$

Then if the thickness of the plate be such that its optical length

for the vibration along OY is half a wave-length greater, or half a wave-length less, than that for the vibration along OX, the displacements at emergence will be

$$u = a \cos \alpha \cos \omega t,$$
$$v = - a \sin \alpha \cos \omega t,$$

so that there is again plane polarization, but the angle of vibration forms an angle $-\alpha$ with the axis of x The same holds for a retardation equal to any odd multiple of two right angles For even multiples, the plane is that of the original vibration These plates, in which a relative retardation of the two waves amounting to half a wave-length takes place, are called "Half Wave-length Plates" and are used in some instruments in which it is desired to fix the plane of polarization accurately The simple Nicol does not permit of very exact adjustment, for while it is moved about near the position of extinction, a broad dark patch is seen to travel across the field, and it is difficult to fix the exact position in which the centre of that patch is in the centre of the field of view In the instrument in which a half-wave plate is used, that plate covers half the field of view If ON and OM, Fig 155, be the principal directions of the half-wave plate covering the left-hand portion of the field of view, and if the incident light vibrates along OP_1, the field of view will be divided by the plate into two portions, the directions of vibration at emergence being along OP_1, OP_2, equally inclined to ON. An eye examining the field through an analysing Nicol will find the two halves unequally illuminated, except where its principal plane is coincident with ON or at right angles to it In the latter position, the luminosity of the field is small if α is small, and the eye is then very sensitive to small differences of illumination, so that the position of the analysing Nicol may be fixed with great accuracy. A half wave-length plate used in this fashion is the distinguishing feature of "Laurent's Polarimeter" The weak point of the arrangement lies in the effect of refrangibility on the retardation, in consequence of which a retardation of half a wave-length can only be obtained for a very limited part of the spectrum Hence homogeneous light must be used with instruments which contain these plates

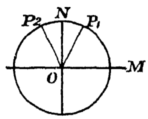

Fig 155

121. The Quarter Wave Plate Plates in which the relative retardation of two waves is a quarter of a period, are called Quarter Wave Plates They have the property of converting plane polarized light vibrating in a suitable direction into circularly polarized light Let OX and OY be the two directions of vibration in the crystal the vibration along OY being the one propagated most quickly

Consider an incident plane polarized ray vibrating at an angle α to OX. The displacements in the incident vibrations are

$$u = a\cos\alpha\cos\omega t,$$
$$v = a\sin\alpha\cos\omega t$$

At emergence the displacements may, by suitable adjustment of the origin of time, be expressed as

$$\left.\begin{aligned} u &= a\cos\alpha\cos\omega t \\ v &= a\sin\alpha\cos(\omega t+\delta) \end{aligned}\right\} \qquad (11)$$

In general this represents an elliptic vibration and we may investigate whether a point P moves clockwise or counter-clockwise through the ellipse. If α is in the first quadrant, then for $\omega t = \pi/2$, the x component of the displacement is zero, and the velocity in the x direction negative. Under these conditions, the rotation is positive (anti-clockwise) or negative (clockwise) according as the y displacement is positive or negative.

But under the above conditions at emergence for $\omega t = \pi/2$

$$v = -a\sin\alpha\sin\delta$$

The rotation is positive or negative, therefore, according as $\sin\delta$ is negative or positive, hence if the total retardation is less than half a wave-length, the rotation is negative or clockwise. We should have got the opposite result if we had taken α to be in the second quadrant. Our conclusions may be formulated thus —

If the retardation is less than half a wave-length, the rotation is from the direction OY, which belongs to the more quickly travelling wave, to the direction OP of the incident vibration, taking that branch of OP which forms an angle less than a right angle with OY.

If the retardation is between half a wave-length and a whole wave-length, the rotation is from the direction OP to the direction OY.

The displacements indicated by (11) when resolved along OP and at right angles to it, become

$$u' = a\left[\cos^2\alpha\cos\omega t + \sin^2\alpha\cos(\omega t+\delta)\right],$$
$$v' = a\left[\sin\alpha\cos\alpha\cos(\omega t+\delta) - \sin\alpha\cos\alpha\cos\omega t\right],$$

and if $\alpha = \dfrac{\pi}{4}$

$$u' = \tfrac{1}{2}a\left[\cos(\omega t+\delta) + \cos(nt)\right] = a\cos\tfrac{1}{2}\delta\cos(\omega t + \tfrac{1}{2}\delta),$$
$$v' = \tfrac{1}{2}a\left[\cos(\omega t+\delta) - \cos nt\right] = -a\sin\tfrac{1}{2}\delta\sin(\omega t+\tfrac{1}{2}\delta),$$

$$\frac{u'^2}{\cos^2\frac{1}{2}\delta} + \frac{v'^2}{\sin^2\frac{1}{2}\delta} = a^2$$

Hence the particle describes in general an ellipse having OP as one of its principal axes When $\delta = \frac{\pi}{2}$ the ellipse becomes a circle If therefore a plane wave be propagated through a doubly refracting substance, and if the incident vibration is equally inclined to OX and OY, then along the normal to the wave, the rays are plane, elliptically, or circularly polarized in regular succession The state of vibration and the direction of rotation are indicated in Fig 156 for equal distances from each other, each step in distance corresponding to a retardation of 45°

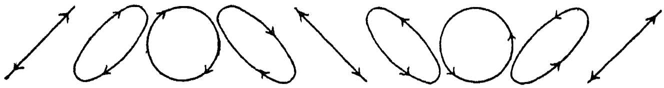

Fig 156

Thin plates of mica or gypsum may be obtained of the right thickness to give circular polarization If the retardation is $3\lambda/4$ the effect is the same, but the rotation is left-handed in the same position relative to the crystal, where it was right-handed with a retardation of $\lambda/4$ We may call the direction OX the *axis* of the quarter plate, so that the direction of rotation for retardation of less than half a wave-length is from the direction of incident vibration to the direction of the axis, through the acute angle included between them A retardation $\lambda/4 + n\lambda$ acts, so far as a particular wave-length is concerned, exactly like one of $\lambda/4$, but the difference in the refractive index for different colours has a more serious effect, the higher the value of n

122. Application of Quarter Wave Plate. Besides being able to give, at any rate for one wave-length, light which is circularly polarized and rotating either in one direction or in the other, a quarter wave plate is useful for the investigation of elliptically polarized light. Elliptic polarization may always be represented by the superposition of two plane vibrations taking place in the direction of the axes of the ellipse and having a relative retardation of 90°. This phase difference is in one direction or another according as the elliptic path is right-handed or left-handed A quarter wave plate with its axis parallel to one of the axes of the ellipse will increase or diminish the existing phase difference by another right angle, and the result is therefore plane polarization If a and b are the semiaxes of the original ellipse, the direction of vibration after passing through the quarter wave plate will form an angle $\tan^{-1}\left(\pm\frac{b}{a}\right)$ with the direction along which a is

measured, the ± sign being determined by the question whether the quarter plate increases or diminishes the original retardation

123. Babinet's Compensator. This is an arrangement which has been successfully used for the study of elliptic polarization It consists of two wedges of quartz, with their axes in the direction of the shading of the two surfaces in Figure 157 If a parallel beam of light traverses the system in the direction LN, the ray vibrating in the direction of the edge CD of the upper prism will pass through that upper prism more quickly, but through the lower prism more slowly, than the vibration at right angles to it The central ray passes through equal thicknesses of both prisms If plane polarized light which may be resolved parallel to AB and AC fall on the prism, the central ray will be plane polarized in the same direction as the incident light but on either side the rays will in general show elliptic polarization At certain distances however the relative retardation of the two rays is two right angles and the transmitted ray will again be plane polarized If the transmitted light be examined by a Nicol, properly placed, the field of view is seen to be traversed by parallel bands If now the original light is elliptically polarized, the whole system of bands is the same as before but shifted sideways. In Babinet's Compensator, each of the wedges may be shifted parallel to itself, and in this way the central band may be brought back to its former position The amount of displacement necessary to bring it back measures the relative retardation, and by its means the ratio of the axes of the ellipse may be determined

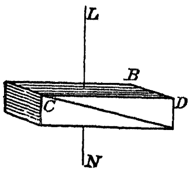

Fig 157

124. Circularly polarized light incident on a crystalline plate. We now consider the case where circularly polarized light falls on a crystalline plate and is then analysed by a Nicol prism or other plane polarizer The incident light may be considered to be made of the superposition of two plane polarized waves having a relative retardation of a quarter of a wave-length To fix our ideas, let the rotation of the incident light be anti-clockwise, the displacement along OX being represented by $a \cos \omega t$ and that along OY by $a \sin \omega t$. The direction of these axes may be chosen according to convenience and we may take them to be coincident with the principal directions of vibration inside the crystal Let there be a retardation δ inside the crystal of that component which vibrates along OY. If the analyser is placed so that the light it can transmit vibrates along a direction forming an angle α with OX (the rotation from OX to OY being positive) the two parts of the beam leaving the analyser have

amplitudes $a \cos \alpha$ and $a \sin \alpha$ and a phase difference of $\frac{1}{2}\pi + \delta$ Hence the intensity of the emergent light is

$$I = a^2 (1 - \sin 2\alpha \sin \delta) \qquad (12)$$

This expression replaces equation (1) which holds when the incident light is plane polarized The achromatic lines are determined by $\sin 2\alpha = 0$, and are therefore two lines at right angles to each other, parallel and perpendicular respectively to the principal plane of the analyser The isochromatic lines are the curves for which δ is constant If the plate is cut from a uniaxal crystal at right angles to the axis, the isochromatic lines are circles which in adjoining quadrants show complementary effects depending on the change of sign of $\sin 2\alpha$

If the plate be examined by convergent or divergent light, the appearance, for positive values of δ, is that shown in Fig 158, and for negative values of δ in Fig 159 As the chromatic influence on the

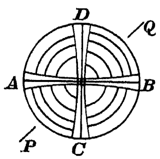

Fig 158

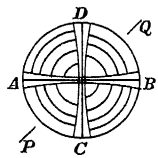

Fig 159

phase difference δ is the greater, the larger the phase difference, the first minimum observed with white light looks darker than the subsequent ones, the minima for the different colours overlapping more closely We may refer to those two minima as the two dark spots, which lie in the first and third quadrants in Fig 158 and in the second and fourth quadrants in Fig 159

The difference in the appearance gives us a useful criterion to distinguish between prolate and oblate crystals Let it be required to study the intensity of light along the line NO (Fig 144), which we take to be the axis of X If OF be the axis of Y at O, and BC the direction of vibration transmitted by the analyser, α is in the first quadrant and $\sin 2\alpha$ in (12) is a positive quantity. OF being the direction of vibration of the ordinary ray, the retardation δ is positive for oblate crystals such as Iceland Spar, in which the ordinary ray is transmitted more slowly Hence Fig 158 represents the appearance If the polarizer is placed at right angles to the analyser or along AB (Fig 144), the axis of the quarter plate must, according to Art 121, be placed parallel to AC if the rotation is to be anti-clockwise, as has been assumed Hence

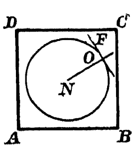

Fig 144.

for oblate crystals, the line forming the dark spots is parallel to the axis of the quarter wave plate, while for prolate crystals (Fig 159) the two lines are at right angles to each other In both figures the line PQ marks the position of the axis of the quarter wave plate It is easily seen that no difference is made in the appearance if the positions of the analyser and polarizer be interchanged Hence the same rule holds whether the original rotation is clockwise or anti-clockwise, and we need only consider the relative positions of the axis of the quarter wave plate and the line joining the dark spots to decide between the two possible kinds of uniaxal crystals

PART II.

CHAPTER X

THEORIES OF LIGHT

125. Small strains in a small volume may always be treated as homogeneous strains Let α, β, γ represent the displacements within a strained body, and let the displacements be expressible as functions of the unstrained coordinates x, y, z of any point, so that

$$\alpha = f_1(x, y, z), \quad \beta = f_2(x, y, z), \quad \gamma = f_3(x, y, z)$$

Let further α', β', γ' be the displacements of a particle near x, y, z, which originally has coordinates $x+\xi$, $y+\eta$ and $z+\zeta$, then, neglecting squares of ξ, η, ζ, by Taylor's theorem

$$\alpha' = \alpha + \frac{d\alpha}{dx}\xi + \frac{d\alpha}{dy}\eta + \frac{d\alpha}{dz}\zeta,$$

$$\beta' = \beta + \frac{d\beta}{dx}\xi + \frac{d\beta}{dy}\eta + \frac{d\beta}{dz}\zeta,$$

$$\gamma' = \gamma + \frac{d\gamma}{dx}\xi + \frac{d\gamma}{dy}\eta + \frac{d\gamma}{dz}\zeta,$$

ξ, η, ζ denoting the coordinates of the second particle relative to those of the first in the unstrained condition. If ξ', η', ζ' denote similarly the relative coordinates of the same two particles in the strained condition, we have

$$\xi' = (x + \xi + \alpha') - (x + \alpha) = \xi + \alpha' - \alpha,$$

or Similarly

$$\left.\begin{aligned} \xi' &= \left(1 + \frac{d\alpha}{dx}\right)\xi + \frac{d\alpha}{dy}\eta + \frac{d\alpha}{dz}\zeta \\ \eta' &= \frac{d\beta}{dx}\xi + \left(1 + \frac{d\beta}{dy}\right)\eta + \frac{d\beta}{dz}\zeta \\ \zeta' &= \frac{d\gamma}{dx}\xi + \frac{d\gamma}{dy}\eta + \left(1 + \frac{d\gamma}{dz}\right)\zeta \end{aligned}\right\} \quad (1).$$

These equations denote a homogeneous strain, for the linear relations between the strained and unstrained coordinates necessarily satisfy all the conditions laid down for such a strain by Thomson and Tait (*Natural Philosophy*, Vol I Art 155) "*If when matter occupying any space is strained in any way, all pairs of points of its substance which are initially at equal distances from one another in parallel lines remain equidistant, it may be at an altered distance; and in parallel lines, altered it may be, from their initial direction; the strain is said to be homogeneous*"

126. Simple Elongation. As a simple example of a homogeneous strain we may take the special case in which all coefficients except $\frac{d\alpha}{dx}$ vanish This gives

$$\xi' = \left(1 + \frac{d\alpha}{dx}\right), \quad \eta' = \eta, \quad \zeta' = \zeta.$$

This is at once seen to represent a strain in which all lines parallel to OX are increased in the ratio $\left(1 + \frac{d\alpha}{dx}\right) : 1$, their distances from each other being unaltered It is therefore a simple elongation along OX, the elongation being measured by $\frac{d\alpha}{dx}$ If $\frac{d\alpha}{dx}$ is small and if $\frac{d\beta}{dy}$ and $\frac{d\gamma}{dz}$ also have values which though small are not negligible, the strain consists of three small elongations along the three coordinate axes, superposed on each other We denote these elongations by e, f, g, so that

$$e = \frac{d\alpha}{dx}, \quad f = \frac{d\beta}{dy}, \quad g = \frac{d\gamma}{dz}$$

A cube having unit sides parallel to the coordinate axes, takes by the strain a volume equal to $(1+e)$ $(1+f)$ $(1+g)$, and neglecting small quantities of the second order, it is seen that the cubical dilatation which is the increase of volume of unit volume is measured by

$$e + f + g = \frac{d\alpha}{dx} + \frac{d\beta}{dy} + \frac{d\gamma}{dz} \quad \ldots \quad (2)$$

In a homogeneous strain all portions of a body have their volume increased or diminished in the same ratio, and we may therefore speak of the dilatation as a quantity belonging to the strain and independent of the position or shape of the portion of the body which we contemplate. This may formally be proved as follows

Take three points having coordinates $\xi_1, \eta_1, \zeta_1, \xi_2, \eta_2, \zeta_2, \xi_3, \eta_3, \zeta_3$ respectively.

The volume τ of the tetrahedron having these points as three of its

rtices and the origin as the fourth, is equal to the sixth part of the terminant

$$\begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix}.$$

This determinant is changed by the strain to

$$\begin{vmatrix} \xi_1' & \eta_1' & \zeta_1' \\ \xi_2' & \eta_2' & \zeta_2' \\ \xi_3' & \eta_3' & \zeta_3' \end{vmatrix}$$

Substituting from equations (1) and applying a well-known theorem determinants, it is found that the volume τ' of the strained tetraedron is the sixth part of the product

$$\begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix} \times \begin{vmatrix} 1+\dfrac{d\alpha}{dx} & \dfrac{d\alpha}{dy} & \dfrac{d\alpha}{dz} \\ \dfrac{d\beta}{dx} & 1+\dfrac{d\beta}{dy} & \dfrac{d\beta}{dz} \\ \dfrac{d\gamma}{dx} & \dfrac{d\gamma}{dy} & 1+\dfrac{d\gamma}{dz} \end{vmatrix}$$

he second determinant simplifies, when the differential coefficients are small that squares may be neglected, and becomes

$$1+\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}$$

Hence
$$\tau'=\tau\left(1+\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}\right),$$

nd
$$\frac{\tau'-\tau}{\tau}=\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}$$

easures the cubical dilatation

127. Simple Shear. Consider a strain which is represented by he equations

$$\xi'=\xi+\frac{d\alpha}{dy}\eta,$$
$$\eta'=\frac{d\beta}{dx}\xi+\eta,$$
$$\zeta'=\zeta$$

A point P on OX (Fig 160), the axis along which both x and ξ are easured, keeps its x coordinate unchanged but is shifted parallel to OY through a distance $\xi d\beta/dx$, so that the line OX is turned through n angle $d\beta/dx$ Similarly a point Q on OY is shifted parallel to OX nd the line OY is turned through an angle $d\alpha/dy$ The parallelogram $OP'RQ'$ has an area which, neglecting small quantities of the second

order, is equal to $OP \times OQ$, so that the strain involves no sensible change of area, and as all z coordinates are unaltered, the strain involves no sensible change of volume. If the strained figure be rotated until OX' coincides with OX, it is seen that the total change may be represented as a sliding of all lines parallel to OX along themselves, the amount of the relative sliding being proportional to the distance between any two lines. The distance QQ' being $OQ\,d\alpha/dy$, is increased by the rotation through an angle $d\beta/dx$ (bringing OP' into coincidence with OP), by an amount $OQ\,d\beta/dx$, so that the sliding per unit distance is

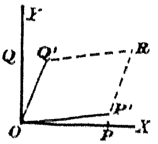

Fig. 160

$$c = \frac{d\alpha}{dy} + \frac{d\beta}{dx}$$

If the total strain is confined to such a sliding, it satisfies the condition of a simple shear (Thomson and Tait, § 171), c being the amount of the shear.

A simple shear may be produced by an elongation e in one direction, together with an equal contraction in a direction at right angles. Let OX and OY (Fig. 161) be the two directions. A length OA is changed by the strain to OA', where $OA' = (1+e)\,OA$. Take a point B on OY at a distance $OB = OA'$. If all lines along OY are reduced in the ratio $(1+e):1$, OB will be changed to OB', so that $OB' = OA$. If $OD = OA$, and $OC = OB$, the parallelogram $ABCD$ will be changed into $A'B'C'D'$. Imagine $A'B'C'D'$ to be transposed so that $A'B'$ is made to coincide with AB, and it will be seen that the whole change is equivalent to a sliding of the lines parallel to AB along their own lengths. If θ be the angle between AD and a line drawn at right angles to AB, the amount of sliding per unit distance is $2\tan\theta$.

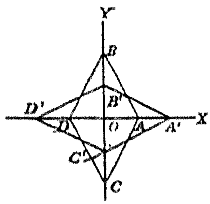

Fig. 161.

If further, α is the angle between OB and AB, $\theta + 2\alpha = \frac{1}{2}\pi$, so that the amount of sliding is $2\cot 2\alpha = \cot\alpha - \tan\alpha$.

Now
$$\tan\alpha = \frac{OA}{OB} = \frac{OA}{OA'} = \frac{1}{1+e}$$
$= 1 - e$ (approximately, if e is a small quantity).

Hence the amount of sliding is $2e$, neglecting small quantities of the second order.

To sum up: "A simple extension in one set of parallels, and a simple contraction of equal amount in any other set perpendicular to those, is the same as a simple shear in either of the two sets of planes cutting the two sets of parallels at 45°. And the numerical

measure of this shear, or simple distortion, is equal to *double* the amount of the elongation or contraction (each measured of course per unit length)" (Thomson and Tait, § 681)

128 Components of Strain Neglecting small quantities of the second order, the strain represented by the equations (1) may be imagined to be produced by the superposition of six separate steps, which are three simple elongations and three simple shears Beginning at first with the three elongations, the resulting change is represented by

$$\xi_1' = \left(1 + \frac{d\alpha}{dx}\right)\xi,$$

$$\eta_1' = \left(1 + \frac{d\beta}{dy}\right)\eta,$$

$$\zeta_1' = \left(1 + \frac{d\gamma}{dz}\right)\zeta$$

We next suppose a change indicated by

$$\xi_2' = \xi_1' + \frac{d\alpha}{dy}\eta_1',$$

$$\eta_2' = \frac{d\beta}{dx}\xi_1' + \eta_1',$$

$$\zeta_2' = \zeta_1',$$

which according to the previous article is a simple shear of amount $c = \frac{d\alpha}{dy} + \frac{d\beta}{dx}$ in the plane of xy By substitution we find, neglecting squares of small quantities, the total change so far to be given by

$$\xi_2' = \left(1 + \frac{d\alpha}{dx}\right)\xi + \frac{d\alpha}{dy}\eta,$$

$$\eta_2' = \frac{d\beta}{dx}\xi + \left(1 + \frac{d\beta}{dy}\right)\eta,$$

$$\zeta_2' = \left(1 + \frac{d\gamma}{dz}\right)\zeta.$$

If we further superpose shears of amount

$$a = \frac{d\beta}{dz} + \frac{d\gamma}{dy} \text{ in the plane of } yz$$

and

$$b = \frac{d\gamma}{dx} + \frac{d\alpha}{dz} \text{ in the plane of } zx,$$

we return to the set of equations (1) The six quantities e, f, g, a, b, c, are called the components of the strain

129. Homogeneous Stress. "When through any space in a body under the action of force, the mutual force between the portions

of matter on the two sides of any plane area is equal and parallel to the mutual force across any equal, similar, and parallel plane area, the stress is said to be homogeneous through that space In other words, the stress experienced by the matter is homogeneous through any space if all equal similar and similarly turned portions of matter within this space are similarly and equally influenced by force " (Thomson and Tait, § 659)

Consider a unit cube (Fig 162) subject to homogeneous internal stresses and in equilibrium The stress on each of the six sides may be decomposed into three along the coordinate axes, but as, from the definition of a homogeneous stress, the forces acting in the same direction across opposite faces must be equal, we need only consider three faces of the cube We denote by X_x, Y_x, Z_x, the three components of force acting on the face yz, the index x indicating that the face is normal to the axis of x Similarly X_y, Y_y, Z_y, and X_z, Y_z, Z_z, indicate the components acting on the faces normal to the axes of y and z respectively If we consider the force which acts on the cube from the outside, two stresses X_x act in opposite directions on the two faces normal to OX If we take X_x to be positive the two forces tend to produce elongation Similarly Y_y and Z_z are stresses tending to produce elongations along the axes of y and z respectively.

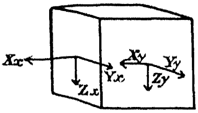

Fig 162

The force X_z (Fig 163) is a tangential force acting in opposite directions on two opposite faces, but not along the same line, so that a couple of moment X_z is formed We take X_z to be positive when, as drawn in the figure, the force acting on a surface parallel to xy from below is along the negative axis of y, the axis of z being positive upwards But the two forces Z_x also form a couple, which however tends to produce rotation round OY in the opposite direction, hence for equilibrium, it follows that

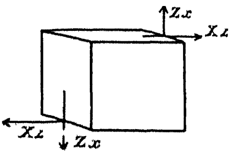

Fig 163

$$X_z = Z_x$$

The two equal couples X_z and Z_x form together a simple shearing stress. It may be proved in the same manner that

$$Y_x = X_y,$$
$$Z_y = Y_z.$$

The six quantities

$$X_x,\ Y_y,\ Z_z,\ Y_z = Z_y,\ Z_x = X_z,\ X_y = Y_x,$$

completely define a homogeneous stress We shall introduce the

notation of Thomson and Tait, and write for these six components of stress

$$P, Q, R, S, T, U$$

130. Shearing stress produced by combined tension and pressure at right angles. Let $ABCD$ be a section of a cube, which is subject to a uniform tension P at right angles to BC, and a uniform pressure at right angles to CD. No stress is supposed to act at right angles to the plane of the paper Let H, K, L, M be the middle points of the sides of the square $ABCD$, and draw the square $HKLM$ If the part HBK is in equilibrium, a force must act on the plane which is at right angles to the plane of the paper, and passes through HK The elementary laws of Statics show that this force must be in the plane, and that its value per unit surface is P The rectangular volume of $HKLM$ is therefore acted on by tangential stresses of the nature of shearing stresses, or

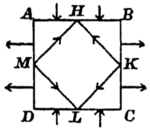

Fig 164

"A longitudinal traction (or negative pressure) parallel to one line and an equal longitudinal positive pressure parallel to any line at right angles to it, is equivalent to a shearing stress of tangential tractions parallel to the planes which cut those lines at 45° And the numerical pressure of this shearing stress, being the amount of the tangential traction in either set of planes, is equal to the amount of the positive or negative normal pressure, *not* doubled" (Thomson and Tait, § 681) The caution at the end of the quotation is necessitated by the fact that in the analogous proposition referring to shears, the amount of the shear is obtained by doubling the elongation, as has been proved in Art 127

131 Connexion between Strains and Stresses. If a simple shearing stress, as defined in Art 129, act on a homogeneous body, it produces a shearing strain, and the ratio of the stress to the strain is the resistance to change of shape or the "Rigidity" of the substance Calling the rigidity n, it follows that we may put

$$S = na\,,\quad T = nb\,,\quad U = nc \qquad (3)$$

in isotropic bodies

The three stresses P, Q, R produce elongations e, f, g, and there must be a linear relationship between them Also by symmetry a stress along OX must produce the same contraction in all directions at right angles to itself Hence A and B being constants, we may write down at once the equations

$$\left.\begin{aligned} P &= Ae + B\,(f+g) \\ Q &= Af + B\,(g+e) \\ R &= Ag + B\,(e+f) \end{aligned}\right\} \qquad (4)$$

It remains to prove how A and B are connected with the rigidity and the bulk modulus. If e, f, g are equal

$$P = Q = R = e(A + 2B)$$

Hence the stress is uniform.

But the cubical dilatation being $3e$ and the bulk modulus being equal to the ratio of the uniform stress P to the cubical dilatation, it follows that

$$3k = A + 2B \qquad (5)$$

As a second special case take $R = 0$, and $Q = -P$, which conditions indicate a shearing stress in planes equally inclined to the axis of X and Y, and these will cause a shearing strain equal in amount to P/n. This shearing strain is equivalent by Art 127 to an elongation in the direction of P of $P/2n$, and an equal contraction in the direction of Q. Substituting $e = -f = P/2n$ into the first of the equations (4), we find if $g = 0$

$$2n = A - B \qquad (6)$$

Combining (5) and (6), it follows that

$$A = k + \tfrac{4}{3}n, \quad B = k - \tfrac{2}{3}n$$

In place of the components of strain, we may introduce their equivalents in terms of the displacement (Arts 126 and 127). Equations (3) and (4) then become

$$S = n\left(\frac{d\beta}{dz} + \frac{d\gamma}{dy}\right), \quad T = n\left(\frac{d\gamma}{dx} + \frac{d\alpha}{dz}\right), \quad U = n\left(\frac{d\alpha}{dy} + \frac{d\beta}{dx}\right) \qquad (7),$$

and

$$\left.\begin{aligned} P &= A\frac{d\alpha}{dx} + B\left(\frac{d\beta}{dy} + \frac{d\gamma}{dz}\right) \\ Q &= A\frac{d\beta}{dy} + B\left(\frac{d\gamma}{dz} + \frac{d\alpha}{dx}\right) \\ R &= A\frac{d\gamma}{dz} + B\left(\frac{d\alpha}{dx} + \frac{d\beta}{dy}\right) \end{aligned}\right\} \qquad (8)$$

132. Equations of Motion in a disturbed medium. Returning to the stresses acting on the cube in Art 129, we consider the case where these stresses are not constant through the volume, but alter slowly from place to place. If the distance between the two faces of the cube which are at right angles to the axis of X is dx, there will be a force

$$X_x dy dz$$

acting in the negative direction on the face which is coincident with the coordinate plane and a force on the opposite face equal to

$$\left(X_x + \frac{dX_x}{dx}dx\right) dy dz.$$

These combine to a resultant

$$\frac{dX_x}{dx}\,dx\,dy\,dz$$

Similarly the force $X_z dx\,dy$ acting on the plane xy in the direction of x together with the force

$$\left(X_z + \frac{dX_z}{dz}\,dz\right)dx\,dy$$

combine to a resultant

$$\frac{dX_z}{dz}\,dx\,dy\,dz,$$

and the forces in that same direction are complete when we have added the resultant

$$\frac{dX_y}{dy}\,dx\,dy\,dz$$

of the two forces which act on the faces which are normal to the axis of y. If ρ be the density of the substance, so that $\rho\,dx\,dy\,dz$ be the mass of the volume considered, and if α be the displacement in the x direction, the equations of motion may be written down by the laws of dynamics, leaving out the factor $dx\,dy\,dz$ on both sides,

$$\rho\frac{d^2\alpha}{dt^2} = \frac{dX_x}{dx} + \frac{dX_y}{dy} + \frac{dX_z}{dz}$$

Similarly

$$\rho\frac{d^2\beta}{dt^2} = \frac{dY_x}{dx} + \frac{dY_y}{dy} + \frac{dY_z}{dz},$$

$$\rho\frac{d^2\gamma}{dt^2} = \frac{dZ_x}{dx} + \frac{dZ_y}{dy} + \frac{dZ_z}{dz}$$

Re-introducing the notation of Thomson and Tait, the equations become

$$\rho\frac{d^2\alpha}{dt^2} = \frac{dP}{dx} + \frac{dU}{dy} + \frac{dT}{dz},$$

$$\rho\frac{d^2\beta}{dt^2} = \frac{dU}{dx} + \frac{dQ}{dy} + \frac{dS}{dz},$$

$$\rho\frac{d^2\gamma}{dt^2} = \frac{dT}{dx} + \frac{dS}{dy} + \frac{dR}{dz}$$

To eliminate the stresses use equations (7) and (8)

Substituting the values of A and B from Art 131, and re-arranging the terms, we obtain

$$\left.\begin{aligned}
\rho\frac{d^2\alpha}{dt^2} &= n\left(\frac{d^2\alpha}{dx^2} + \frac{d^2\alpha}{dy^2} + \frac{d^2\alpha}{dz^2}\right) + (k + \tfrac{1}{3}n)\frac{d}{dx}\left(\frac{d\alpha}{dx} + \frac{d\beta}{dy} + \frac{d\gamma}{dz}\right)\\
\rho\frac{d^2\beta}{dt^2} &= n\left(\frac{d^2\beta}{dx^2} + \frac{d^2\beta}{dy^2} + \frac{d^2\beta}{dz^2}\right) + (k + \tfrac{1}{3}n)\frac{d}{dy}\left(\frac{d\alpha}{dx} + \frac{d\beta}{dy} + \frac{d\gamma}{dz}\right)\\
\rho\frac{d^2\gamma}{dt^2} &= n\left(\frac{d^2\gamma}{dx^2} + \frac{d^2\gamma}{dy^2} + \frac{d^2\gamma}{dz^2}\right) + (k + \tfrac{1}{3}n)\frac{d}{dz}\left(\frac{d\alpha}{dx} + \frac{d\beta}{dy} + \frac{d\gamma}{dz}\right)
\end{aligned}\right\} \quad (9)$$

These equations govern wave propagation in all elastic media We may obtain from them the characteristic equations for the longitudinal waves of sound by putting the rigidity n of the medium equal to zero When applied to light, the medium is taken as incompressible, so that

$$\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}=0,$$

but k at the same time becomes infinitely large Writing

$$\delta=\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz},$$

$$p=k\delta,$$

and

$$\nabla^2=\frac{d^2}{dx^2}+\frac{d^2}{dy^2}+\frac{d^2}{dz^2},$$

the equations become

$$\left.\begin{aligned}\rho\frac{d^2\alpha}{dt^2}&=n\nabla^2\alpha+\frac{dp}{dx}\\ \rho\frac{d^2\beta}{dt^2}&=n\nabla^2\beta+\frac{dp}{dy}\\ \rho\frac{d^2\gamma}{dt^2}&=n\nabla^2\gamma+\frac{dp}{dz}\end{aligned}\right\} \qquad (10)$$

These equations, together with certain relations which must hold at the surfaces of the elastic body, constitute the elastic solid theory of light

For plane waves, the displacements are the same at all points of the wave-front, which we may imagine to be at right angles to the axis of z The differential coefficient of α, β, γ with respect to x and y must therefore vanish The equations (9) then reduce to

$$\rho\frac{d^2\alpha}{dt^2}=n\frac{d^2\alpha}{dz^2},\quad \rho\frac{d^2\beta}{dt^2}=n\frac{d^2\beta}{dz^2},\quad \rho\frac{d^2\gamma}{dt^2}=(k+\tfrac{4}{3}n)\frac{d^2\gamma}{dz^2}. \qquad (11)$$

The last equation represents a longitudinal wave propagated with infinite velocity and having no relation to any observed phenomenon of light Each of the first two equations represents a rectilinear wave propagated with velocity $\sqrt{n/\rho}$, a result already deduced by the simpler but less general methods of Art 12

The investigation of wave propagation in crystalline media presents great difficulties The simplest hypothesis from a mathematical point of view is that of assuming that the inertia of the medium may differ for displacements in different directions By substituting ρ_1, ρ_2, ρ_3, respectively, for ρ on the left-hand side of equations (9), we obtain equations which lead to a wave surface which is similar to, but not identical with, Fresnel's wave surface. A theory of double refraction

based on this hypothesis was brought forward by Lord Rayleigh*, but abandoned because observations made by Stokes, and afterwards by Glazebrook, decided in favour of Fresnel's surface Instead of taking the inertia as variable, we may adopt the very plausible hypothesis that the rigidity is different in different directions Thus different values of n in the first two equations (11) would give two plane waves propagated with different velocities, along the axis of z A general theory cannot however be formed by a simple modification of the equations holding for isotropic media According to Green†, there may be twenty-one different coefficients defining the properties of crystalline media, which shows the complication we might be led into if we wished to attack the problem in its most general form

133 Equations of the Electromagnetic Field The line integral of the magnetic force round a closed curve is numerically equal to the electric current through the curve multiplied by 4π It is shown in treatises on Electricity that the mathematical expression of this law is contained in the three equations

$$\left.\begin{aligned} 4\pi u &= \frac{d\gamma}{dy} - \frac{d\beta}{dz} \\ 4\pi v &= \frac{d\alpha}{dz} - \frac{d\gamma}{dx} \\ 4\pi w &= \frac{d\beta}{dx} - \frac{d\alpha}{dy} \end{aligned}\right\} \qquad (12),$$

where α, β, γ are the components of magnetic force, and u, v, w the components of current density The factor 4π depends on the units chosen, which are those of the electromagnetic system

Another proposition which embodies Faraday's laws of electromagnetic induction states that if a closed curve encloses lines of magnetic induction which vary in intensity, an electromotive force acts round the curve, and the line integral of the electric force round the closed curve is equal to the rate of diminution of the total magnetic induction through the circuit This leads to the equations

$$\left.\begin{aligned} -\mu\frac{d\alpha}{dt} &= \frac{dR}{dy} - \frac{dQ}{dz} \\ -\mu\frac{d\beta}{dt} &= \frac{dP}{dz} - \frac{dR}{dx} \\ -\mu\frac{d\gamma}{dt} &= \frac{dQ}{dx} - \frac{dP}{dy} \end{aligned}\right\} \qquad (13),$$

where $\mu\alpha$, $\mu\beta$, $\mu\gamma$ are the components of magnetic induction, μ being the permeability, and P, Q, R those of electric force.

* *Collected Works*, Vol I p 111

† *Collected Works*, p 245

The two sets of equations may be taken to represent experimental facts and to be quite independent of any theory, although equations (13) may be deduced from (12) with the help of the principle of the conservation of energy Both sets of equations would be equally true if we considered electric and magnetic forces to be due to action at a distance

There are some additional equations to be considered

Differentiating equations (12) with respect to x, y, z respectively and adding, we find

$$\frac{du}{dx} + \frac{dv}{dy} + \frac{dw}{dz} = 0 \tag{14}$$

Similarly we derive from (13), if μ be constant and α, β, γ periodic,

$$\frac{d\alpha}{dx} + \frac{d\beta}{dy} + \frac{d\gamma}{dz} = 0 \tag{15}$$

134. Maxwell's Theory. The fundamental principle of Maxwell lies in his conception of an electric current in dielectrics and the way in which this current is made to depend on electric force His views are best explained by an analogy taken from the theory of stress and strain A stress in an elastic solid produces certain displacements which are proportional to the stress If the stresses increase, the displacements increase, and the change of displacements constitutes a transference of matter This flow or current of matter is proportional to the rate of change of the elastic stress Taking this as a guide we may imagine the medium to yield in some unknown manner to the application of electric force, and if so, the rate of change of that force will be proportional to a "flow" which according to Maxwell is identical in all its effects with an electric current

If the electric force is E, the electric current is proportional to $\frac{dE}{dt}$, and if the law that the total flow is the same across all cross-sections of a circuit holds good for these so-called "displacement currents" or "polarization currents," it can be shown that the current is equal to $K\frac{dE}{dt}\Big/4\pi$, where K is the specific inductive capacity of the medium In a conductor, the current would, according to Ohm's law, be CE, where C is the conductivity If we imagine a medium to possess both specific inductive capacity and conductivity, we must introduce an expression which includes both cases and put the current equal to

$$\left(C + \frac{1}{4\pi} K \frac{d}{dt}\right) E \quad \ldots \ldots \ldots \tag{16}$$

Confining ourselves at present to non-conductors and resolving along the three coordinate axes, we have

$$u=\frac{1}{4\pi}K\frac{dP}{dt}, \quad v=\frac{1}{4\pi}K\frac{dQ}{dt}, \quad w=\frac{1}{4\pi}K\frac{dR}{dt} \qquad (17)$$

These equations allow us to combine (12) and (13) so as to obtain two fresh sets containing respectively only the magnetic and the electric forces

135. Differential equation for propagation of electric and magnetic disturbances in dielectric media.

Equations (12) with the help of (17) become

$$\left.\begin{aligned} K\frac{dP}{dt}&=\frac{d\gamma}{dy}-\frac{d\beta}{dz}\\ K\frac{dQ}{dt}&=\frac{d\alpha}{dz}-\frac{d\gamma}{dx}\\ K\frac{dR}{dt}&=\frac{d\beta}{dx}-\frac{d\alpha}{dy}\end{aligned}\right\} \qquad (18)$$

Differentiate each of the equations (13) with respect to the time, eliminate P, Q, R, by means of (18), and use (15), when the following sets of equations, involving only magnetic forces, will be obtained

$$K\mu\frac{d^2\alpha}{dt^2}=\nabla^2\alpha, \quad K\mu\frac{d^2\beta}{dt^2}=\nabla^2\beta, \quad K\mu=\frac{d^2\gamma}{dt^2}=\nabla^2\gamma \qquad (19).$$

We may eliminate the magnetic forces in a similar manner and obtain

$$K\mu\frac{d^2P}{dt^2}=\nabla^2P, \quad K\mu\frac{d^2Q}{dt^2}=\nabla^2Q, \quad K\mu\frac{d^2R}{dt^2}=\nabla^2R \qquad (20)$$

These equations show that the magnetic and electric forces are propagated with a velocity $1/\sqrt{K\mu}$ In the electromagnetic system of units, $\mu=1$ in vacuo, and differs very little from that value in any known dielectric K the specific inductive capacity is, in vacuo, unity when the *electrostatic* system of units is employed, but in the electromagnetic system K is numerically equal to $1/v^2$, if v is equal to the number of electrostatic units of quantity which are contained in an electromagnetic unit This number, which gives the velocity of propagation of electromagnetic waves in vacuo, may be determined by experiment, and is found, within the errors of experiment, to be equal to the velocity of light in vacuo Both velocities differ from 3×10^{10} probably by not more than one part in a thousand

Maxwell's theory, which is embodied in equations (19) or (20), leads therefore to the remarkable conclusion that an electromagnetic disturbance is propagated with a finite velocity which is equal to the velocity of light This conclusion has been amply verified by the

celebrated experiments of Hertz Kirchhoff* had already in 1857 pointed out that a longitudinal electric disturbance is propagated in a wire with a velocity equal to that of light, but it was left to Maxwell to discover the reason for this coincidence

If both the disturbance of light and the electromagnetic wave are propagated through the same medium with the same velocity, the conclusion is irresistible that both phenomena are identical in character. This conclusion constitutes the so-called "Electromagnetic Theory of Light" The electromagnetic theory of light establishes for the propagation of a luminous disturbance, equations which in several instances, as will appear, fit the facts better than the older elastic solid theory, but it should not be forgotten that it furnishes no explanation of the nature of light It only expresses one unknown quantity (light) in terms of other unknown quantities (magnetic and electric disturbances), but magnetic and electric stresses are capable of experimental investigation, while the elastic properties of the medium through which, according to the older theory, light was propagated, could only be surmised from the supposed analogy with the elastic properties of material media Hence it is not surprising that the electromagnetic equations more correctly represent the actual phenomena Whatever changes be introduced in future, in our ideas of the nature of light, the one great achievement of Maxwell, the proof of the identity of luminous and electromagnetic disturbances, will never be overthrown.

136. Refraction. We have so far only considered the propagation of waves in vacuo According to equations (20), the squares of the velocities of propagation in two media having identical magnetic permeabilities, ought to be inversely as their specific inductive capacities If therefore K_0 be the inductive capacity of the vacuum, K_1 that of any dielectric, the "refractive index" ought to be equal to $\sqrt{K_1/K_0}$ This relation is approximately verified in the case of a few gases, as shown in the following table, which contains the square roots of specific inductive capacities (D) as measured by Klemencic†, and the refractive indices (n) of the same gases for Sodium light, as measured by G. W Walker‡ Both constants are reduced to a temperature of 0° C, and a pressure of 760 mm.

Nature of Gas	(D)	(n)
Air	1 000293	1 000293
Hydrogen	1 000132	1 000141
Carbon dioxide	1 000492	1 000451
Sulphur dioxide	1 000477	1 000676

* *Pogg Ann* Vol c p 193 (1857)

† *Wien Ber* (2) Vol xci p 1 (1885)

‡ *Trans Roy Soc* A Vol cci p 435 (1903)

The discrepancy for sulphur dioxide is already well marked

For solids and liquids the relation altogether fails Thus water has a specific inductive capacity which is 80 times greater than that of air, and its refractive index should therefore be equal to 9, or six times larger than its actual value But these discrepancies are not surprising, for we have left a factor out of consideration, which to a great extent dominates the phenomenon of refraction, and that is absorption. The theoretical relationship really applies only to waves of infinite length, but in most cases we know nothing of the refractive index for very long waves The subject will be further discussed in the next Chapter

137. Direction of Electric and Magnetic Forces at right angles to each other. If we confine ourselves for the sake of simplicity to waves, parallel to the plane of xy, we must take in equations (13) and (17) all quantities to be independent of x and y. these equations then become

$$\left.\begin{aligned} \mu\frac{d\alpha}{dt}&=\frac{dQ}{dz}, & \mu\frac{d\beta}{dt}&=-\frac{dP}{dz}, & \mu\frac{d\gamma}{dt}&=0, \\ K\frac{dP}{dt}&=-\frac{d\beta}{dz}, & K\frac{dQ}{dt}&=\frac{d\alpha}{dz}, & K\frac{dR}{dt}&=0 \end{aligned}\right\} \qquad (21)$$

It follows that there is no component of either the electric or the magnetic force normal to the plane of the wave, and that therefore the whole of the disturbance is in that plane If the electric disturbance is in one direction only, so that *e g* $Q=0$, it follows that $\alpha=0$, or that the magnetic disturbance is also rectilinear, and at *right angles to the electric disturbance* We have therefore for the simplest case of a plane wave, two vectors representing the electric and magnetic forces respectively, and these vectors are at right angles to each other and to the direction of propagation

More generally let the components P and Q of a plane wave-front parallel to xy be

$$P=\phi(z-vt),\quad Q=\psi(z-vt),$$

so that
$$v\frac{dP}{dz}=-\frac{dP}{dt},\quad v\frac{dQ}{dz}=-\frac{dQ}{dt},$$

or making use of (21)

$$\mu v\frac{d\beta}{dt}=\frac{dP}{dt},\quad \mu v\frac{d\alpha}{dt}=-\frac{dQ}{dt}. \qquad (22)$$

We thus find
$$\mu v\beta=P,\quad \mu v\alpha=-Q,$$

and hence
$$\alpha P+\beta Q=0$$

This shows that also in this more general case the electric and magnetic forces are at right angles to each other

138. Double Refraction. In crystalline substances the specific inductive capacity of a plate may depend on the direction in which the plate is cut, relative to the axes of the crystals. The currents which are generated in such substances by a variation of electric force are not necessarily in the direction of the force, but if P, Q, R be the electric forces resolved in three directions at right angles to each other, and if the current in any one direction be a linear function of P, Q, R, then it may be proved that there are always three directions at right angles to each other such that the current is in the direction of the force. If we choose these directions for the coordinate axes, we may write

$$u = \frac{1}{4\pi} K_1 \frac{dP}{dt}, \qquad v = \frac{1}{4\pi} K_2 \frac{dQ}{dt}, \qquad w = \frac{1}{4\pi} K_3 \frac{dR}{dt} \qquad (23),$$

where K_1, K_2, K_3, are the three principal dielectric constants.

These equations replace (16). The elimination of α between (13) and (17) now leads to

$$\left.\begin{aligned} K_1 \mu \frac{d^2P}{dt^2} &= \nabla^2 P - \frac{d}{dx}\left(\frac{dP}{dx} + \frac{dQ}{dy} + \frac{dR}{dz}\right) \\ K_2 \mu \frac{d^2Q}{dt^2} &= \nabla^2 Q - \frac{d}{dy}\left(\frac{dP}{dx} + \frac{dQ}{dy} + \frac{dR}{dz}\right) \\ K_3 \mu \frac{d^2R}{dt^2} &= \nabla^2 R - \frac{d}{dz}\left(\frac{dP}{dx} + \frac{dQ}{dy} + \frac{dR}{dz}\right) \end{aligned}\right\} \qquad (24)$$

If $K_1 = K_2 = K_3$, $\qquad \dfrac{dP}{dx} + \dfrac{dQ}{dy} + \dfrac{dR}{dz} = 0,$

and we are brought back to the equations which have already been deduced for isotropic media. We proceed to investigate under what conditions plane waves are propagated in a medium to which equations (24) apply. If l, m, n, are the direction cosines of the normal of the plane wave, and V the velocity of propagation, all variable quantities must be expressible as functions of $lx + my + nz - Vt$.

We may therefore in the case of a rectilinear disturbance write for P, Q, R,

$$P_0 f(lx + my + nz - Vt), \quad Q_0 f(lx + my + nz - Vt),$$
$$R_0 f(lx + my + nz - Vt) \qquad (25),$$

where P_0, Q_0, R_0 are constants defining the direction of the electric disturbance, the cosines of the angles which the direction of the electric force forms with the coordinate axes being as $P_0 : Q_0 : R_0$.

By substitution, equations (24) become, if we write

$$v_1 = 1/\sqrt{K_1\mu}, \quad v_2 = 1/\sqrt{K_2\mu}, \quad v_3 = 1/\sqrt{K_3\mu}$$

and $$S = lP + mQ + nR\,;$$

$$\left.\begin{aligned} P &= \frac{lSv_1^2}{v_1^2 - V^2} \\ Q &= \frac{mSv_2^2}{v_2^2 - V^2} \\ R &= \frac{nSv_3^2}{v_3^2 - V^2} \end{aligned}\right\} \quad \cdots \quad (26).$$

Multiplying the first of these equations by l, the second by m, and the third by n, and adding, we obtain the characteristic equation for V,

$$\frac{v_1^2 l^2}{v_1^2 - V^2} + \frac{v_2^2 m^2}{v_2^2 - V^2} + \frac{v_3^2 n^2}{v_3^2 - V^2} = 1,$$

or subtracting $$l^2 + m^2 + n^2 = 1,$$

$$\frac{l^2}{v_1^2 - V^2} + \frac{m^2}{v_2^2 - V^2} + \frac{n^2}{v_3^2 - V^2} = 0 \quad \ldots\ldots (27).$$

This is an equation identical with (4) Chapter VIII, and shows that the electromagnetic wave theory leads to the correct construction for the propagation of plane waves

From (26) we also derive

$$\frac{lP}{v_1^2} + \frac{mQ}{v_2^2} + \frac{nR}{v_3^2} = \left(\frac{l^2}{v_1^2 - V^2} + \frac{m^2}{v_2^2 - V^2} + \frac{n^2}{v_3^2 - V^2}\right) S = 0.$$

As P/v_1^2, Q/v_2^2, R/v_3^2, are proportional to the components of electric current, we conclude that the electric current is in the plane of the wave-front

The substitution of (25) into (13) leads to

$$V\mu\alpha = Rm - Qn,$$
$$V\mu\beta = Pn - Rl,$$
$$V\mu\gamma = Ql - Pm\,,$$

from which it follows that

$$l\alpha + m\beta + n\gamma = 0,$$

and $$P\alpha + Q\beta + R\gamma = 0$$

Hence the magnetic force is in the plane of the wave, and the electric force is at right angles to the magnetic force, though not in general, as will presently appear, in the plane of the wave

In Art 84 it was found that if an ellipsoid

$$v_1^2 x^2 + v_2^2 y^2 + v_3^2 z^2 = 1 \quad \ldots\ldots \quad \ldots\ldots\ldots (28)$$

be constructed, the reciprocals of the two principal axes of any plane section measure the two velocities of plane waves which are parallel to the section, and it was proved that this construction leads to equation (27) This equation has been shown to lead to Fresnel's

wave-surface which is therefore now established as a consequence of the electromagnetic theory Propositions with respect to wave or ray velocities which are proved in the same chapter may all be interpreted in terms of the electromagnetic theory if we take the components of the electric current u, v, w to correspond to the displacements in the older theory

If an electric disturbance is propagated as a plane wave, and a normal be drawn to the ellipsoid (28) at the end of the vector having u, v, w as components, the direction cosines of this normal are proportional to $v_1^2 u$, $v_2^2 v$, $v_3^2 w$ and are therefore by (22) coincident with the direction cosines of the vector representing the electric force This electric force is therefore not in the plane of the wave but lies in a plane which contains the wave normal and the electric current. It has been shown in Art 86 that this plane also contains the ray

139. Problem of refraction and reflexion. A good test of the adequacy of any theory of light is found in its capability of dealing with the problem of reflexion and refraction Reflexion takes place when a wave falls on a surface at which the properties of the medium are suddenly changed If the transition is gradual, there is no reflexion A ray of light *e g* enters our atmosphere from outside and gradually passes into denser and denser layers of air Though its path becomes curved by refraction, there is no reflexion, and neglecting absorption, the intensity of the ray remains unaltered The fact that a surface of glass or water partially reflects a ray of light shows that the transition between the media of different refractive indices must take place within a distance not much greater than a wave-length

Before entering into the relative merits of different theories with regard to the problem of reflexion, we may deduce some general results which are independent of any theory We consider a plane wave-front having its normal in the plane xy Its displacements, in whatever direction they are, must be capable of expression in the form $f(ax + by - ct)$, for

$$ax + by - ct = \text{constant}$$

expresses a plane parallel to the axis of z If θ be the angle between the wave normal and the axis of x, and v the velocity of wave propagation, we have

$$\cos\theta = \frac{a}{\sqrt{a^2 + b^2}}, \quad \sin\theta = \frac{b}{\sqrt{a^2 + b^2}}, \quad \cot\theta = \frac{a}{b},$$

and
$$v = \frac{c}{\sqrt{a^2 + b^2}}$$

The line of intersection of the wave with the plane $x=0$ is

$$by - ct = \text{constant}$$

and travels forward therefore parallel to itself with a uniform speed of c/b

If the plane $x=0$ is a surface at which refraction takes place, the displacement in the refracted wave may be expressed as

$$F(a_1 x + b_1 y - c_1 t).$$

If the displacements are periodic, the periods must be the same in the refracted and incident beam, hence $c_1 = c$ Also the lines of intersection of the refracted and incident waves with the surface $x=0$ must always be coincident Hence this line must travel forward with the same velocity in both waves This proves that $b_1 = b$ The velocity of the refracted wave is

$$v_1 = \frac{c_1}{\sqrt{a_1^2 + b_1^2}} = \frac{c}{\sqrt{a_1^2 + b^2}}$$

Hence calling the angle of refraction θ_1

$$\frac{v}{v_1} = \frac{\sqrt{a_1^2 + b^2}}{\sqrt{a^2 + b^2}} = \frac{\sin\theta}{\sin\theta_1}$$

This proves the law of refraction The displacements in the reflected wave will be of the form

$$F(a'x + b'y - c't)$$

The previous reasoning shows that $c' = c$, $b' = b$. Also the velocity of wave propagation must now be identical with that of the incident wave. Hence $a' = \pm a$ We must choose the lower sign, as otherwise the wave would simply return in the original direction The numerical equality of a' and a proves the law of reflexion

140. Reflexion in the Electromagnetic Theory. The problem of reflexion is comparatively simple if treated according to the electromagnetic theory, and we shall therefore begin with it. In the electrostatic or electromagnetic field the electric and magnetic forces have to satisfy certain conditions at the surface of separation of two media having different properties These are in treatises on electricity proved to be the following: (1) The tangential components of *electric force* are the same on both sides of the surface (2) The normal components of *electric displacement* are continuous (3) The tangential components of *magnetic force* are continuous, and (4) the normal components of *magnetic induction* are continuous Taking the surface $x=0$ to be the surface of separation, we may put with the previous notation these so-called surface conditions into the form.

$$K\frac{dP}{dt} = K_1\frac{dP_1}{dt}, \quad Q = Q_1, \quad R = R_1 \qquad \dots (29),$$

$$\mu\alpha = \mu_1\alpha_1, \quad \beta = \beta_1, \quad \gamma = \gamma_1 \qquad \dots (30)$$

The right-hand sides of all equations apply to the second medium which we shall refer to as the *lower* medium, taking the axis x positive downwards

These six equations are not all independent The continuity of Q and R involves the continuity of their variations in any tangential direction and hence the first equation (13) shows that the continuity of normal magnetic induction is secured Similarly the continuity of the tangential components of magnetic force leads to the continuity of the normal electric current as shown from the first of equations (18). We may therefore omit the first of equations (29) and (30) as being contained in the others, nevertheless it is often convenient to introduce them

If the wave-front be parallel to the axis of z

$$dP/dz = 0$$

Also writing without appreciable error $\mu_1 = \mu$ for all transparent media, the continuity of β is satisfied according to (13) if dR/dx is continuous. We may therefore replace the surface conditions by the following five.

$$\left.\begin{aligned} KP = K_1 P_1, \quad Q = Q_1, \quad R = R_1 \\ \frac{dR}{dx} = \frac{dR_1}{dx}, \quad \frac{dQ}{dx} - \frac{dP}{dy} = \frac{dQ_1}{dx} - \frac{dP_1}{dy} \end{aligned}\right\} \quad . \qquad (31)$$

We now take the incident beam to be plane polarized and first treat the case that the electric force is at right angles to the plane of incidence which we take to be the plane of xy Therefore $P = Q = 0$, and the surface conditions reduce to

$$\left.\begin{aligned} R = R_1 \\ \frac{dR}{dx} = \frac{dR_1}{dx} \end{aligned}\right\} \quad . \qquad (32)$$

For the electric force in the incident wave we may write $e^{\iota(ax+by-ct)}$ and for that of the reflected wave $r e^{\iota(-ax+by-ct)}$, where the real parts only need ultimately be retained A change of phase will be indicated by a complex value of r If s is the amplitude of the transmitted beam, we have therefore

$$\left.\begin{aligned} R = e^{\iota(ax+by-ct)} + r e^{\iota(-ax+by-ct)} \\ R_1 = s e^{\iota(a_1 x+by-ct)} \end{aligned}\right\} \qquad (33),$$

in the upper and lower media respectively

The surface conditions give at once for $x = 0$

$$1 + r = s,$$

$$a(1-r) = a_1 s,$$

$$r = \frac{a - a_1}{a + a_1}$$

$$= \frac{\cot\theta - \cot\theta_1}{\cot\theta + \cot\theta_1}$$

$$= \frac{\sin(\theta_1 - \theta)}{\sin(\theta_1 + \theta)} \quad . \quad . \quad . \qquad (34).$$

The square of this expression correctly represents the observed intensity of the reflected beam, if the incident beam is polarized in the plane of incidence We conclude that the electric force is at right angles to the plane of polarization, a result in accordance with the conclusion arrived at in the study of double refraction

If we take the incident beam to be polarized at right angles to the plane of incidence, R, α, β vanish, and the surface conditions become

$$\left.\begin{aligned} KP = K_1P_1, \quad Q = Q_1 \\ \frac{dQ}{dx} - \frac{dP}{dy} = \frac{dQ_1}{dx} - \frac{dP_1}{dy} \end{aligned}\right\} \qquad (35)$$

The last equation secures the continuity of γ But the form of our assumed disturbance shows that $d\gamma/dy = ib\gamma$ and hence if γ is continuous so is also $d\gamma/dy$ and vice versa Also according to (18) $K\,dP/dt = d\gamma/dy$, when $\beta = 0$ the first and last surface conditions are therefore identical and we may disregard the latter

If WF (Fig 165) be the incident wave-front, the displacement is now in the plane of the paper and parallel to the wave-front Let the direction indicated by the arrow be that in which the displacements are taken to be positive $W'F'$ represents the reflected wave-front, and we may again arbitrarily fix that direction for which we shall take the displacements to be positive

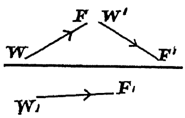

Fig 165

It is obvious that for normal incidence there is no distinction between this case and the one already considered when the displacement is at right angles to the plane of incidence It is therefore convenient to take that direction as positive which agrees with that of the incident wave when the incidence is normal The arrow indicates the direction Similarly for the transmitted wave W_1F_1 Taking the amplitude of the incident beam again to be unit amplitude, and resolving along OX and OY, we may put in the upper medium

$$P = -\sin\theta\, e^{i(ax+by-ct)} + r\sin\theta\, e^{i(-ax+by-ct)},$$

$$Q = \cos\theta\, e^{i(ax+by-ct)} + r\cos\theta\, e^{i(-ax+by-ct)},$$

and in the lower medium

$$P_1 = -s\sin\theta_1\, e^{(a_1x+by-ct)},$$

$$Q_1 = s\cos\theta_1 e^{(a_1x+by-ct)}$$

The condition $KP = K_1P_1$ for $x = 0$ gives

$$\frac{(1-r)}{v^2}\sin\theta = \frac{s\sin\theta_1}{v_1^{\,2}},$$

or

$$(1-r)\sin\theta_1 = s\sin\theta,$$

and the condition $Q = Q_1$ gives

$$(1+r)\cos\theta = s\cos\theta_1$$

These are the only conditions that need be satisfied.

Eliminating s we obtain

$$(1-r)\sin\theta_1\cos\theta_1=(1+r)\sin\theta\cos\theta,$$

or

$$r=\frac{\sin\theta_1\cos\theta_1-\sin\theta\cos\theta}{\sin\theta_1\cos\theta_1+\sin\theta\cos\theta}$$

$$=\frac{\sin 2\theta_1-\sin 2\theta}{\sin 2\theta_1+\sin 2\theta}$$

$$=\frac{\tan(\theta_1-\theta)}{\tan(\theta_1+\theta)} \quad \ldots\ldots\ldots(36).$$

This again is a formula agreeing with observation, at any rate as a first approximation The application of the equations (34) and (36) to the cases of oblique polarization or unpolarized light has already been discussed in Art. 27 as well as the observed departures from (36).

It has often been suggested that the experimental deviations from the tangent law may be due to the fact that the transition between the two media is not sudden but takes place within a layer comparable in thickness with the length of a wave. L Lorenz* first investigated the question and showed that a thickness of from the tenth to the hundredth part of a wave-length is sufficient to cause the observed effect. Drude†, treating the same subject from the standpoint of the electromagnetic theory, has arrived at similar results, a thickness of the transition layer of ·0175 λ being found to be sufficient in the case of flint-glass to account for the elliptic polarization observed near the polarizing angle

141. Reflexion in the elastic solid theory. In elastic solids the conditions at the boundary are obtained by the consideration that as a tearing of the medium can only take place under application of forces which exceed the limits of elasticity, the displacements on both sides of the boundary must be the same, while the medium is performing oscillations under the conditions of perfect elasticity.

A second condition is imposed by the third law of motion. *The stresses must be continuous* The continuity of stress together with that of displacement satisfies also the requirements of the law of conservation of energy, as the work done across any surface is the product of stress and rate of change of displacement

The components of displacement which we had previously called α, β, γ, shall, in order to distinguish them from the magnetic forces for which we have introduced the same letters, now be designated by ξ, η, ζ

* *Pogg Ann* CXI. p 460 (1860) and CXIV p 238 (1861)

† *Lehrbuch der Optik*, p. 266

The continuity of displacement introduces the conditions

$$\xi=\xi_1,\quad \eta=\eta_1,\quad \zeta=\zeta_1,$$

where the right-hand sides refer to the lower medium

The stresses on a surface normal to the axis of x are, by Art 131,

$$P=A\frac{d\xi}{dx}+B\left(\frac{d\eta}{dy}+\frac{d\zeta}{dz}\right),$$

$$T=n\left(\frac{d\zeta}{dx}+\frac{d\xi}{dz}\right),\quad U=n\left(\frac{d\xi}{dy}+\frac{d\eta}{dx}\right),$$

where $\qquad A=k+\frac{4}{3}n,\quad B=k-\frac{2}{3}n$

Writing $m=k+\frac{1}{3}n$, we obtain for the conditions of continuity of stress

$$\left.\begin{aligned}(m+n)\frac{d\xi}{dx}+(m-n)\left(\frac{d\eta}{dy}+\frac{d\zeta}{dz}\right)&\\ =(m_1+n_1)\frac{d\xi_1}{dx}+(m_1-n_1)\left(\frac{d\eta_1}{dy}+\frac{d\zeta_1}{dz}\right)&\\ n\left(\frac{d\zeta}{dx}+\frac{d\xi}{dz}\right)=n_1\left(\frac{d\zeta_1}{dx}+\frac{d\xi_1}{dz}\right)&\\ n\left(\frac{d\xi}{dy}+\frac{d\eta}{dx}\right)=n_1\left(\frac{d\xi_1}{dy}+\frac{d\eta_1}{dx}\right)&\end{aligned}\right\}\quad \ldots\ (37),$$

where m_1, n_1, define the elastic properties of the second medium.

Let the plane of xy be the plane of incidence, and the vibrations of a plane wave be at right angles to that plane All displacements vanish except ζ, and ζ is independent of z Hence the equations of continuity reduce to

$$\zeta=\zeta_1,\quad n\frac{d\zeta}{dx}=n_1\frac{d\zeta_1}{dx}.$$

The equation of motion in the upper medium is, according to (9),

$$\frac{d^2\zeta}{dt^2}=\frac{n}{\rho}\left(\frac{d^2\zeta}{dx^2}+\frac{d^2\zeta}{dy^2}\right),$$

with a similar equation for the lower medium

But $\qquad \zeta=e^{\iota(ax+by-ct)}+re^{\iota(-ax+by-ct)},$

$$\zeta_1=se^{\iota(a_1x+by-ct)}.$$

For $x=0$, the surface conditions give

$$1+r=s,$$

$$na(1-r)=n_1a_1s,$$

and eliminating s,

$$r=\frac{na-n_1a_1}{na+n_1a_1}=\frac{n\cot\theta-n_1\cot\theta_1}{n\cot\theta+n_1\cot\theta_1}\quad \ldots.\ (38)$$

For the velocity of wave propagation in an elastic solid, we have $v^2=n/\rho$ Different wave velocities in different media may either be

due to differences in the rigidity or to differences in density Hence we must distinguish the two cases.

Case I $n = n_1$.

Equation (33) becomes

$$r = \frac{\cot\theta - \cot\theta_1}{\cot\theta + \cot\theta_1}$$
$$= \frac{\sin(\theta_1 - \theta)}{\sin(\theta + \theta_1)}$$

This agrees with the result obtained in the electromagnetic theory if the displacements are made to correspond to electric force

Case II

$$\rho = \rho_1, \quad \frac{n_1}{n} = \frac{v_1^2}{v^2} = \frac{\sin^2\theta_1}{\sin^2\theta}$$

Equation (33) now gives

$$r = \frac{\sin^2\theta\tan\theta_1 - \sin^2\theta_1\tan\theta}{\sin^2\theta\tan\theta_1 + \sin^2\theta_1\tan\theta}$$
$$= \frac{\sin 2\theta - \sin 2\theta_1}{\sin 2\theta + \sin 2\theta_1}$$
$$= \frac{\tan(\theta - \theta_1)}{\tan(\theta + \theta_1)}$$

This is the equation for the reflected light when the incident wave is polarized at right angles to the plane of incidence Hence if different media differ by their rigidities, the reflexion of light vibrating at right angles to the plane of incidence can only be accounted for by supposing that the plane of polarization contains the vibration.

To work out completely the more complicated case that the vibration lies in the plane of incidence, we must transform the equations of motion

In equations (9) alter the notation, put $\zeta = 0$, and let ξ and η be independent of z, this being the condition that the wave normal lies in the plane of xy The equations then become

$$\rho\frac{d^2\xi}{dt^2} = m\frac{d}{dx}\left(\frac{d\xi}{dx} + \frac{d\eta}{dy}\right) + n\left(\frac{d^2\xi}{dx^2} + \frac{d^2\xi}{dy^2}\right),$$

or

$$\left.\begin{aligned}\rho\frac{d^2\xi}{dt^2} &= (m+n)\frac{d}{dx}\left(\frac{d\xi}{dx} + \frac{d\eta}{dy}\right) + n\frac{d}{dy}\left(\frac{d\xi}{dy} - \frac{d\eta}{dx}\right)\\ \text{Similarly}\quad \rho\frac{d^2\eta}{dt^2} &= (m+n)\frac{d}{dy}\left(\frac{d\xi}{dx} + \frac{d\eta}{dy}\right) - n\frac{d}{dx}\left(\frac{d\xi}{dy} - \frac{d\eta}{dx}\right)\end{aligned}\right\} \quad (39)$$

Introducing two new functions such that

$$\xi = \frac{d\phi}{dx} + \frac{d\psi}{dy}, \quad \eta = \frac{d\phi}{dy} - \frac{d\psi}{dx} \quad . \quad . \quad . \quad (40),$$

we find that (39) may be satisfied by

$$\left.\begin{aligned} \rho \frac{d^2\phi}{dt^2} &= (m+n)\left(\frac{d^2\phi}{dx^2} + \frac{d^2\phi}{dy^2}\right) \\ \text{and} \qquad \rho \frac{d^2\psi}{dt^2} &= n\left(\frac{d^2\psi}{dx^2} + \frac{d^2\psi}{dy^2}\right) \end{aligned}\right\} \quad . \qquad (41)$$

From (35) it appears that the displacements ξ and η, due to changes of ϕ, are at right angles to the surface $\phi =$ constant, while the displacements due to changes in ψ lie *in* the surface $\psi =$ constant. If we adopt the same form of solution for ϕ and ψ, it is the latter function which gives the motion which we require for the propagation of light in which the displacements are in the wave-front We put therefore for the incident wave $\psi = e^{i(ax+by-ct)}$, and assume for the form of solution generally,

In the upper medium

$$\left.\begin{aligned} \psi &= e^{i(ax+by-ct)} + re^{i(-ax+by-ct)} \\ \phi &= pe^{i(a'x+by-ct)} \end{aligned}\right\} \quad .. \qquad (42).$$

In the lower medium

$$\left.\begin{aligned} \psi_1 &= se^{i(a_1x+by-ct)} \\ \phi_1 &= qe^{i(a_1'x+by-ct)} \end{aligned}\right\} \quad .. \qquad (43).$$

Substituting these values in equations (41) we obtain

$$c^2 = \frac{n}{\rho}(a^2+b^2) = \frac{n_1}{\rho_1}(a_1^2+b^2) = \frac{(m+n)}{\rho}(a'^2+b^2)$$
$$= \frac{m_1+n_1}{\rho_1}(a_1'^2+b^2) \qquad (44).$$

From the first two equalities we obtain as before the law of refraction, but as m and m_1 are indefinitely great the last equalities give

$$a'^2+b^2=0\,, \quad a_1'^2+b^2=0 \quad \ldots \quad . \qquad (45)$$

For the same reason

$$\frac{d^2\phi}{dx^2}+\frac{d^2\phi}{dy^2}=\frac{d^2\phi_1}{dx^2}+\frac{d^2\phi_1}{dy^2}=0 \quad . \quad . \quad . \qquad (46).$$

This shows that the motion due to ϕ is that of an incompressible liquid. As ϕ represents the velocity potential, the motion is irrotational. Also by substitution of (45) into (42) and (43) retaining only the real parts.

$$\phi = pe^{-bx}\cos(by-ct),$$
$$\phi_1 = qe^{-bx}\cos(by-ct)$$

The displacements in so far as they are due to ϕ are

$$\frac{d\phi}{dx} = -bpe^{-bx}\cos(by-ct) \qquad \text{(upper medium)},$$

$$\frac{d\phi_1}{dx} = -bqe^{-bx}\cos(by-ct) \qquad \text{(lower medium)},$$

and $$\frac{d\phi}{dy} = -bpe^{-bx}\sin(by-ct) \qquad \text{(upper medium)},$$

$$\frac{d\phi_1}{dy} = -bqe^{-bx}\sin(by-ct) \qquad \text{(lower medium)}$$

The motion vanishes for normal incidence as $b = \frac{2\pi}{\lambda}\sin\theta$ Unless θ is small, the exponential factor shows that the motion quickly diminishes with the distance from the refracting surface

The surface conditions are

$$\xi = \xi_1\,; \quad \eta = \eta_1\,,$$

$$(m+n)\frac{d\xi}{dx} + (m-n)\frac{d\eta}{dy} = (m_1+n_1)\frac{d\xi_1}{dx} + (m_1-n_1)\frac{d\eta_1}{dy},$$

or in terms of ψ and ϕ

$$(m+n)\frac{d^2\phi}{dx^2} + (m-n)\frac{d^2\phi}{dy^2} + 2n\frac{d^2\psi}{dx\,dy} = \text{similar expressions},$$

or introducing (46)

$$n\left\{2\frac{d^2\psi}{dx\,dy} + \frac{d^2\phi}{dx^2} - \frac{d^2\phi}{dy^2}\right\} = \text{similar expressions}.$$

The quantities p, q, r, s may now be obtained by substituting ϕ, ψ, ϕ_1, ψ_1 from equations (42) and (43) Green*, to whom the above investigation is due, assumes $n = n_1$, and Rayleigh† has put the solution for r in that case into the form

$$r^2 = \frac{\cot^2(\theta+\theta') + M^2}{\cot^2(\theta-\theta') + M^2},$$

where $M = \frac{\mu^2-1}{\mu^2+1}$ and μ is the refractive index

If two media do not differ much in optical properties, so that the refractive index is nearly equal to one, we obtain for the ratio of amplitudes the expression

$$\frac{\tan(\theta-\theta')}{\tan(\theta+\theta')},$$

as required by experiment when the vibration takes place at right angles to the plane of incidence

As has been pointed out above, the tangent formula is only approximately correct, but the deviations are not so great as those which Green's formula would lead us to expect, and are sometimes in the other direction

The alternative according to which differences in optical properties are due to differences in elasticity, leads to results which can in no way be reconciled with observed facts If we place ourselves on the standpoint of the elastic solid theory, we are therefore compelled to conclude

* *Collected Works*, p 245. † *Collected Works*, vol. I p. 129.

that the rigidity of the æther is the same in all media. Even then we arrive at an unsatisfactory result so far as light polarized at right angles to the plane of incidence is concerned

142. Lord Kelvin's theory of contractile æther. According to the most general equations of the motion of an elastic substance (Art 132), a disturbance spreads in the form of two waves, the condensational longitudinal wave propagated with a velocity $\sqrt{(k+\frac{4}{3}n)/\rho}$ and the transverse distortional wave propagated with a velocity $\sqrt{n/\rho}$ The phenomena of light leave no room for a longitudinal wave propagated with finite velocity It has been got rid of in the theory so far considered by taking the elastic body as incompressible The coefficient k then is infinitely large, and the longitudinal disturbance is propagated with infinite velocity

This elastic solid theory of the æther, as discussed in the preceding investigations, does not, however, consistently lead to facts which are in agreement with observation It fails to account for the laws of double refraction and for the observed amplitude of light reflected from transparent bodies That theory was therefore considered dead, until Lord Kelvin* resuscitated it in a different form by showing how, dropping the hypothesis of "solidity," an elastic theory of the æther may still be a possible one

The characteristic distinction of the new theory lies in the bold assumption that the velocity of the longitudinal wave, instead of being infinitely large, is infinitely small This requires that $k+\frac{4}{3}n$ shall be zero, so that k is negative A medium in which there is a negative resistance to compression would at first sight appear to be essentially unstable, but Lord Kelvin shows that the instability cannot come into play, if the æther is rigidly attached to a bounding surface So long as there is a finite propagational velocity for each of the two kinds of wave motion, any disturbance set up in the medium cannot lead to instability Putting therefore the constant A of Article 131 equal to zero, and taking the rigidity to be equal in all media, Lord Kelvin has shown that the theory leads to Fresnel's tangent formula for the amplitude of light polarized in a plane perpendicular to the plane of incidence. Glazebrook† then showed that the consideration of double refraction leads to Fresnel's wave surface, while J. Willard Gibbs‡ pointed out that the new form of elastic æther theory must always lead to the same equation as the electromagnetic theory, provided we replace the symbol which denotes 'displacement' in one theory by that which denotes 'force' in the other and vice versa

* *Phil. Mag* xxvi p 414 1888
† *Phil Mag* xxvi p 521 1888
‡ *Phil Mag* xxvii p 238 1889.

If in equations (9) we write $k = -\frac{4}{3}n$, and allow different values of ρ according as the displacements are in the direction x, y, or z, the equations become with our present notation

$$\left.\begin{aligned}\frac{\rho_1}{n}\frac{d^2\xi}{dt^2} &= \nabla^2\xi - \frac{d}{dx}\left(\frac{d\xi}{dx} + \frac{d\eta}{dy} + \frac{d\zeta}{dz}\right)\\ \frac{\rho_2}{n}\frac{d^2\eta}{dt^2} &= \nabla^2\eta - \frac{d}{dy}\left(\frac{d\xi}{dx} + \frac{d\eta}{dy} + \frac{d\zeta}{dz}\right)\\ \frac{\rho_3}{n}\frac{d^2\zeta}{dt^2} &= \nabla^2\zeta - \frac{d}{dz}\left(\frac{d\xi}{dx} + \frac{d\eta}{dy} + \frac{d\zeta}{dz}\right)\end{aligned}\right\} \quad \ldots \quad . (47).$$

These equations are identical with (24) provided that we replace P, Q, R in the latter by ξ, η, ζ, and μ, K_1, K_2, K_3 by $1/n$, ρ_1, ρ_2, ρ_3, respectively As regards surface conditions, we must now remember that the resistance to compression being negative, there may be infinite compression or dilatation at any point or surface at which a condensational wave tends to start The surface at which reflexion takes place gives rise according to the preceding article to condensational waves, hence disregarding this wave which can only be propagated with zero velocity, the conditions which hold in the general elastic theory, in which the condensational wave is considered, are not necessarily satisfied They must be replaced by others, which, as J W Gibbs has shown, may be obtained directly from the equations of motion

Introduce new quantities defined by

$$\xi' = \frac{d\zeta}{dy} - \frac{d\eta}{dz}, \quad \eta' = \frac{d\xi}{dz} - \frac{d\zeta}{dx}, \quad \zeta' = \frac{d\eta}{dx} - \frac{d\xi}{dy},$$

$$\xi'' = \frac{d\zeta'}{dy} - \frac{d\eta'}{dz}, \quad \eta'' = \frac{d\xi'}{dz} - \frac{d\zeta'}{dx}, \quad \zeta'' = \frac{d\eta'}{dx} - \frac{d\xi'}{dy}$$

Performing the differentiation so as to obtain ξ'', η'', ζ'' in terms of ξ, η, ζ, we arrive at the expressions which stand on the right-hand side of (47) with the sign reversed We take the boundary to be normal to the axis of x If η' and ζ' were discontinuous at such a boundary, $\frac{d\eta'}{dx}$ and $\frac{d\zeta'}{dx}$ would be infinitely large, and this would make η'' and ζ'', and consequently also $\frac{d^2\eta}{dt^2}$, $\frac{d^2\zeta}{dt^2}$ infinitely large, which is obviously not admissible Hence we conclude that η' and ζ' are continuous A similar reasoning shews that as η' and ζ' must remain finite, η and ζ must be continuous The conditions of continuity are therefore

$$\eta = \eta_1, \quad \zeta = \zeta_1,$$

$$\frac{d\xi}{dz} - \frac{d\zeta}{dx} = \frac{d\xi_1}{dz} - \frac{d\zeta_1}{dx}, \quad \frac{d\eta}{dx} - \frac{d\xi}{dy} = \frac{d\eta_1}{dx} - \frac{d\xi_1}{dy}.$$

But these are exactly the conditions which we have seen must hold in the electromagnetic theory if ξ, η, ζ are replaced by P, Q, R The analogy is made complete if it is noticed that the continuity of ξ' and ξ'' follows from the above equations, and that according to the first of equations (47) if $v_1 = \sqrt{n}/\rho$,

$$\frac{1}{v_1^2}\frac{d^2\xi}{dt^2} = -\xi''$$

The continuity of ξ'' is seen to involve the continuity of the normal force and this corresponds to the continuity of electric displacement in the electromagnetic theory.

The analogy has therefore been proved both for the equations regulating the motion of the medium and for the surface conditions.

If this analogy is kept in view, all the results which have been found to hold in the electromagnetic theory may be translated at once into consequences of the contractile æther theory. Thus in the theory of double refraction, the displacements of the latter theory are not in the wave surface, but are normal to the ray as has been shown for the electric forces in Art 133 Thus adopting the contractile æther theory we may conclude at once that when plane waves are propagated through a doubly refracting medium the elastic force and not the displacement is in the plane of the wave I have given a statement of this theory on account of its mathematical interest, but it has now been abandoned by its author*

143. Historical. AUGUSTIN LOUIS CAUCHY, born August 21st, 1789, in Paris, died May 23rd, 1857, at Sceaux, near Paris, was one of the large number of celebrated French mathematicians who, during the end of the 18th and the beginning of the 19th century, made the first serious advance in Mathematical Physics since Newton's time Cauchy's contribution to the theory of light consisted in initiating the endeavour to deduce the differential equations for the motion of light from a theory of elasticity This theory was based on definite assumptions of the actions between the ultimate particles of matter The luminiferous æther like other matter was supposed to be made up of distinct centres of force acting upon each other according to some law depending on the distance. Cauchy explained the phenomena of dispersion by supposing that in the media in which dispersion takes place, the distance between the ultimate particles is no longer small compared with the wave-length He thus arrived at a formula which for a long time was considered to represent satisfactorily the connexion between wave-length and refractive index (Art 150) Cauchy also showed that metallic reflexion may be accounted for by a high

* *Baltimore Lectures*, p. 214.

coefficient of absorption. by interpreting Fresnel's sine and tangent formula, in the case where the index of refraction is imaginary, he obtained the equations for the elliptic polarization of light reflected from metallic surfaces, which are still adopted as correctly representing the facts

GEORGE GREEN was born at Sneinton, near Nottingham, in 1793, and only entered the University of Cambridge at the age of 40 Having graduated in 1837 as fourth wrangler, he was elected to a fellowship in Gonville and Caius College in 1839, and died in 1841 The following paragraph which stands at the head of his celebrated Memoir on the Reflexion and Refraction of Light will show the ideas which guided him in his work

"M Cauchy seems to have been the first who saw fully the utility of applying to the Theory of Light those formulæ which represent the motions of a system of molecules acting on each other by mutually attractive and repulsive forces, supposing always that in the mutual action of any two particles, the particles may be regarded as points animated by forces directed along the right line which joins them This last supposition, if applied to those compound particles, at least, which are separable by mechanical division, seems rather restrictive, as many phenomena, those of crystallisation for instance, seem to indicate certain polarities in these particles. If, however, this were not the case, we are so perfectly ignorant of the mode of action of the elements of the luminiferous ether on each other, that it would seem a safer method to take some general physical principle as the basis of our reasoning, rather than assume certain modes of action, which, after all, may be widely different from the mechanism employed by nature, more especially if this principle include in itself, as a particular case, those before used by M Cauchy and others, and also lead to a much more simple process of calculation. The principle selected as the basis of the reasoning contained in the following paper is this · In whatever way the elements of any material system may act upon each other, if all the internal forces exerted be multiplied by the elements of their respective directions, the total sum for any assigned portion of the mass will always be the exact differential of some function But, this function being known, we can immediately apply the general method given in the *Mécanique Analytique*, and which appears to be more especially applicable to problems that relate to the motions of systems composed of an immense number of particles mutually acting upon each other One of the advantages of this method, of great importance, is, that we are necessarily led by the mere process of the calculation, and with little care on our part, to all the equations and conditions which are

requisite and sufficient for the complete solution of any problem to which it may be applied."

The function introduced above by Green we now call "Potential Energy," and a particular interest attaches to the whole paper, as it is the first instance of the application of the principle of Conservation of Energy to a great physical problem. Green shows that in the most general case, there may be twenty-one different coefficients defining the elastic properties of a medium, and that these reduce to two in the case of an isotropic or uncrystallized medium. The conditions which hold at the surface of two media are deduced, and for the first time strict dynamic principles were applied to the calculation of the amplitudes of the reflected and refracted light. Assuming the difference in the optical behaviour of different media to be differences of density, the Fresnel sine formula is obtained for light polarized in the plane of incidence, and it is shown that for light polarized at right angles to the plane of incidence, the tangent formula can only hold approximately. In a further paper "On the propagation of light in crystallized media," we meet the difficulties which have so long beset all attempts to account satisfactorily for Fresnel's wave surfaces, and though this paper will still be read with advantage, its interest at present is only historical.

GEORGE GABRIEL STOKES, born August 13th, 1819, at Screen in Ireland, graduated as Senior Wrangler in 1841, and was elected to the Lucasian Chair of Mathematics in Cambridge in 1849. He died on February 1st, 1903. His celebrated Memoir "On the Dynamical Theory of Diffraction" contains the complete solution of the problem of the propagation of waves through an elastic medium. The question is treated in so masterly a manner that though published in the year 1849, the paper should still be carefully studied by every student of Optics. He published other important optical memoirs, of which the following may specially be quoted: "On the theory of certain bands seen in the spectrum" (1848), "On the formation of the Central Spot of Newton's Rings beyond the critical angle" (1848), "On Hardinger's Brushes" (1850), "Report on Double Refraction" (1862).

Many of his writings on the theory of sound and hydrodynamics have also optical applications. Stokes was the first to recognize the true nature of fluorescence, only isolated facts as to the luminescence of certain substances under the action of light having been previously known. He made a thorough experimental investigation which proved the possibility of a change in the refrangibility of light (*Phil. Trans.* 1852 and 1853).

JAMES CLERK MAXWELL, born June 13th, 1831, at Edinburgh, died at Cambridge November 5th, 1879, was the first occupant of the

Cavendish Professorship of Physics at Cambridge, which he held from the date of its foundation in 1871 to the time of his death He was one of the most original minds who ever turned their attention to scientific enquiry All mathematicians who, previous to Maxwell, had discussed the undulatory theory of Optics, started from the elastic solid theory of the luminiferous æther That theory was able to give a satisfactory account of a great number of the phenomena of light and was considered to be securely established The phenomena of electricity were treated as independent facts, though no doubt many physicists held that ultimately electric action would be explained by the stresses and strains of the same medium which transmitted light No one had, however, suggested properties of the medium different from those of an ordinary elastic solid Maxwell attacked the question with great originality from another point of view. Having asked himself the question, what the properties of a medium must be, in order that it should be capable of transmitting electric actions, he discovered that this electric medium was capable of transmitting transverse vibrations with the velocity of light Maxwell also showed how Fresnel's wave-surface in double refracting media could be obtained by assuming that, in such media, there may be three dielectric constants, the polarization measured along three axes at right angles to each other being different Of his other optical writings, his memoir "On the theory of Compound Colours and the relations of the Colours of the Spectrum" (*Phil Trans* 1860) deserves special mention

CHAPTER XI.

DISPERSION AND ABSORPTION

144. Wave-fronts with varying amplitudes. We have hitherto confined our attention to vibrations having the same amplitude along each wave-front In other words, the surfaces of equal phase were coincident with the surfaces of equal amplitude. We shall now treat the question in a more general manner, starting from the differential equation of the wave propagation in an absorbing medium which, as we shall see, may be put into the form

$$G\frac{d^2R}{dt^2}+F\frac{dR}{dt}=\frac{d^2R}{dx^2}+\frac{d^2R}{dy^2}+\frac{d^2R}{dz^2} \quad . \quad . \quad . \quad (1),$$

where G and F are constants and R represents the displacement in the elastic solid theory or the electric force in the electromagnetic theory measured in the z direction If the wave-front is plane and parallel to the axis of z, R is independent of z If the disturbance is simply periodic, so that the time only occurs in the form of a periodic factor $pe^{-\iota\omega t}$, where p may be real or imaginary or complex, equation (1) is equivalent to

$$-(\iota F\omega+G\omega^2)R=\frac{d^2R}{dx^2}+\frac{d^2R}{dy^2},$$

or

$$\frac{d^2R}{dx^2}+\frac{d^2R}{dy^2}+\vartheta^2R=0 \quad . \quad . \quad . \quad . \quad . \quad (2),$$

where ϑ^2 depends only on the frequency $\omega/2\pi$ of the disturbance; ϑ^2 is real when F is zero and only in that case. For a particular solution of (2) and therefore of (1), we have

$$R=R_0e^{\iota(ax+by-\omega t)} \quad . \quad . \quad . \quad (3)$$

The substitution of this value of R into (2) leads to the condition

$$a^2+b^2=\vartheta^2 \quad . \quad . \quad . \quad . \quad . \quad (4),$$

showing that a^2+b^2 must be independent of the *direction* in which the wave is propagated. If a and b are both real, there is no absorption and we may put

$$a=\frac{2\pi}{\lambda'}\cos\theta, \quad b=\frac{2\pi}{\lambda'}\sin\theta; \quad \omega=\frac{2\pi v}{\lambda'},$$

where λ' is the wave-length in the medium. The planes of equal phase are represented by

$$x\cos\theta + y\sin\theta = \text{constant}.$$

Let, in the more general case, a and b be complex and write therefore

$$a = \frac{2\pi}{\lambda'}\cos\theta + k_1 i,$$

$$b = \frac{2\pi}{\lambda'}\sin\theta + k_2 i$$

Then (3) may be written, retaining the real parts,

$$R = R_0 e^{-(k_1 x + k_2 y)} \cos\frac{2\pi}{\lambda'}(x\cos\theta + y\sin\theta - vt) \qquad (5).$$

The amplitude is the same over planes satisfying the equation

$$k_1 x + k_2 y = \text{constant},$$

but these planes do not necessarily coincide with the planes of equal phase We have, however, still to satisfy the condition (4) It will be convenient here for the sake of obtaining symmetrical expressions to write

$$k_1 = \frac{2\pi}{\lambda}\kappa\cos\alpha,$$

$$k_2 = \frac{2\pi}{\lambda}\kappa\sin\alpha,$$

where λ is in vacuo the wave-length corresponding to a given ω, *i e* $\omega = 2\pi V/\lambda$ (V being the velocity of light in vacuo). We also write

$$\frac{V}{v} = \frac{\lambda}{\lambda'} = \nu \quad \dots \qquad (6)$$

Hence
$$a = \frac{2\pi}{\lambda}(\nu\cos\theta + i\kappa\cos\alpha),$$

$$b = \frac{2\pi}{\lambda}(\nu\sin\theta + i\kappa\sin\alpha),$$

$$a^2 + b^2 = \frac{4\pi^2}{\lambda^2}\{\nu^2 - \kappa^2 + 2i\nu\kappa\cos(\theta - \alpha)\} \quad \dots \qquad (7)$$

In transparent media $a^2 + b^2$ is real, hence either $\kappa = 0$ or $\cos(\theta - \alpha) = 0$ The first alternative leads to the case already discussed, of waves of equal amplitude But the second alternative shows the possibility of waves of unequal amplitude being transmitted as plane waves, provided that the surfaces of equal amplitude are at right angles to the surfaces of equal phase If we take the axis of x for the direction of propagation, (5) takes the form

$$R = R_0 e^{-\frac{2\pi}{\lambda}\kappa y} \cos\frac{2\pi\nu}{\lambda}(x - vt) \quad \dots \qquad (8)$$

One important and somewhat unexpected result follows. *The velocity with which the wave-front proceeds depends on the variation of amplitude along it* This is shown by (7) for

$$\vartheta^2 = a^2 + b^2 = \frac{4\pi^2}{\lambda^2}(\nu^2 - \kappa^2) \qquad (9),$$

$$\therefore \nu^2 = \frac{\lambda^2\vartheta^2}{4\pi^2} + \kappa^2 = \nu_1^2 + \kappa^2,$$

if ν_1 be the particular value which ν takes when κ is zero. It now follows from (6) that for v the velocity of the wave-front we have

$$v = \frac{V}{\sqrt{\nu_1^2 + \kappa^2}}$$

The velocity with which a disturbance is transmitted is represented by the *ray* velocity V/ν_1 which must of course always be the same in the same medium But our investigation is important as showing that even in vacuo the ray need not be at right angles to the wave-front, and that if this is the case the velocity of the wave-front is not the velocity with which a disturbance starting from a point is transmitted. If ϕ be the angle between the ray and the wave normal

$$\cos\phi = \frac{\nu_1 v}{V} = \frac{\nu_1}{\sqrt{\nu_1^2 + \kappa^2}}$$

In order that these effects should be observable, the change in amplitude must be very rapid, for, according to (8), a change in y equal to e g $10^4\lambda$ (about 5 cm) would, if $\kappa = \frac{1}{2\pi}10^{-4}$, be accompanied by a diminution of the amplitude in the ratio of e 1 Even this rapid change in amplitude would not cause any observable change in the velocity of propagation which depends on the square of κ, the ray would be inclined to the normal at an angle of about 3 5 seconds of arc. It is only when we come to deal with the transmission of waves in highly absorbent substances that our theoretical deductions find their observable applications When absorption takes place ϑ^2 is no longer real Substituting for that quantity its value in terms of F and G, and combining (4) and (7), we find

$$\left.\begin{aligned} \nu^2 - \kappa^2 &= GV^2 \\ 2\nu\kappa\cos\rho &= \frac{\lambda}{2\pi}FV \end{aligned}\right\} \qquad (10),$$

where ρ represents the angle between the planes of equal phase and the planes of equal amplitude

Both ν and κ now depend on the angle ρ, and I shall write ν_0, κ_0 for the values of ν, κ in the particular case that $\rho = 0$

When a plane wave falls normally on an absorbing medium, all parts of the transmitted wave-front have passed through the same

thickness of the medium Hence in that case, the wave-front is also a surface of equal amplitude and $\rho = 0$. If the wave-front is normal to the axis of x, so that $\sin\theta = \sin\alpha = 0$, (5) becomes

$$R = R_0 e^{-\frac{2\pi}{\lambda}\kappa_0 x} \cos 2\pi \frac{\nu_0}{\lambda}(x - vt) \tag{11}$$

As ν_0, κ_0 may be calculated from (10), and in this case $v = V/\nu_0$, all quantities are determined κ_0 is called the coefficient of "extinction."

When the light falls obliquely on the surface the planes of equal amplitude remain parallel to the surface but the planes of equal phase become inclined to the surface at an angle ρ which is the angle of refraction The expression for the disturbance becomes

$$R = R_0 e^{-\frac{2\pi}{\lambda}\kappa x} \cos 2\pi \frac{\nu}{\lambda}(x\cos\rho + y\sin\rho - vt)$$

Both κ and ν depend on the angle ρ, but (10) shows that $\nu^2 - \kappa^2$ and $\nu\kappa\cos\rho$ are constant, which gives the relations

$$\left.\begin{aligned} \nu^2 - \kappa^2 &= \nu_0^2 - \kappa_0^2 \\ \nu\kappa\cos\rho &= \nu_0\kappa_0 \end{aligned}\right\} \quad . \quad . \quad (12)$$

Ketteler, who first realized the importance of these equations, called them the principal equations of wave propagation in absorbing media. The optical distance between two points on the same wave normal, at a distance d apart, is $\frac{2\pi}{\lambda}\nu d$, and we may call ν the coefficient of optical length

We obtain from (10)

$$\left.\begin{aligned} \nu_0^2 - \kappa_0^2 &= GV^2 \\ 2\nu_0\kappa_0 &= FV^2/\omega \end{aligned}\right\} \tag{13}$$

145. The laws of refraction in absorbing media. Let the disturbance of the incident light be proportional to

$$e^{\iota(ax+by-\omega t)},$$

the refracted disturbance will be proportional to

$$e^{\iota(a_1x+by-\omega t)},$$

the identity of the coefficients b and ω on the two sides of the separating surface being proved, as in the case of transparent media. If λ/ν is the wave-length of the transmitted light and ρ the angle of refraction

$$\sin\theta = \nu\sin\rho,$$

or

$$\nu^2\cos^2\rho = \nu^2 - \sin^2\theta$$

Equations (12) may now be written

$$\nu^2 - \kappa^2 = \nu_0^2 - \kappa_0^2,$$

$$\nu^2\kappa^2 - \kappa^2\sin^2\theta = \nu_0^2\kappa_0^2.$$

Solving these we find

$$\left.\begin{aligned} 2\nu^2 &= \sqrt{(\nu_0^2-\kappa_0^2-\sin^2\theta)+4\nu_0^2\kappa_0^2}+(\nu_0^2-\kappa_0^2+\sin^2\theta) \\ 2\kappa^2 &= \sqrt{(\nu_0^2-\kappa_0^2-\sin^2\theta)^2+4\nu_0^2\kappa_0^2}-(\nu_0^2-\kappa_0^2-\sin^2\theta) \end{aligned}\right\} \quad (14)$$

These equations represent Ketteler's law of refraction for absorbing media Having found ν for a particular incidence θ by means of the first of equations (14), $\rho = \sin^{-1}(\nu^{-1}\sin\theta)$ gives the angle of refraction

In the case of transparent media, the wave velocities in the two media are $c/\sqrt{a^2+b^2}$ and $c/\sqrt{a_1^2+b^2}$ respectively, and the refractive index is

$$\mu = \sqrt{\frac{a_1^2+b^2}{a^2+b^2}} \quad \ldots \quad (15)$$

We shall take this equation as a definition of the symbol μ also in the case of opaque media Though μ is now a complex quantity and has no physical meaning, it is useful as an intermediate variable.

Let the first medium be transparent and the vacuum for convenience, so that $a^2+b^2=4\pi^2/\lambda^2$ Applying equation (7) to the second medium, and noting that $\theta-\alpha$ is the angle we now denote by ρ, we then find

$$\begin{aligned} \mu^2 = \frac{a_1^2+b^2}{a^2+b^2} &= (\nu^2-\kappa^2)+2i\kappa\nu\cos\rho \\ &= \nu_0^2-\kappa_0^2+2i\kappa_0\nu_0, \end{aligned}$$

$$\therefore \quad \mu = \nu_0 + i\kappa_0 \quad \ldots\ldots \quad (16)$$

The constants ν_0 and κ_0 are therefore derived directly from μ, being respectively its real and imaginary part

We associate with μ an unreal angle θ_1 defined by

$$\sin\theta = \mu\sin\theta_1 .$$

From this we obtain

$$\begin{aligned} \mu^2\cos^2\theta_1 &= \mu^2-\sin^2\theta \\ &= \nu^2-\kappa^2+2i\kappa\nu\cos\rho-\sin^2\theta \\ &= \nu^2\cos^2\rho-\kappa^2+2i\kappa\nu\cos\rho \\ &= (\nu\cos\rho+i\kappa)^2, \end{aligned}$$

$$\therefore \quad \sin\theta\cot\theta_1 = m + i\kappa \quad \ldots\ldots \quad (17),$$

where $$m = \nu\cos\rho \quad \ldots\ldots \quad (18)$$

146. Free and Forced Vibrations A particle attracted to a fixed centre by a force varying as the distance, was proved in Chapter I to perform simple periodic oscillations. We may extend the investigation now, by admitting the possibility that the motion of the

particle is resisted by some frictional force which acts in proportion to its velocity The path being rectilinear, the equation of motion is

$$\ddot{z} + 2k\dot{z} + n^2 z = 0,$$

having for solution

$$z = Ae^{-kt} \cos \{\sqrt{n^2 - k^2}\ t - \alpha\} \quad . \qquad (19),$$

where A and α are two constants of integration to be determined by the initial conditions of motion It will be noticed that the period of oscillation is increased by the friction and that each maximum of displacement is smaller than its predecessor The motion is no longer simply periodic, and could not give rise to homogeneous waves

If the friction be so great that $n^2 - k^2$ is negative, the form of the solution alters and the motion becomes "aperiodic," but for our present purpose, we may leave this case out of account When κ is small, the period may be expressed in terms of a series

$$\frac{2\pi}{\sqrt{n^2 - k^2}} = 2\pi \left[\frac{1}{n} + \frac{k^2}{2n^3} + \quad \right],$$

which shows that the most important term involving k depends on its square, so that even though we may take account of effects depending on the first power of k, the period is not affected by friction if we may neglect the second power Equation (19) represents the motion which the particle assumes when unacted on by external forces, and is therefore called the *free* vibration

Let the same particle be now subject to an additional periodic force of period $2\pi/\omega$ Its equation of motion becomes

$$\ddot{z} + 2k\dot{z} + n^2 z = E \cos \omega t,$$

where, m being the mass, $E \cos \omega t/m$ is the force The complete solution now is

$$z = \frac{E \sin \epsilon}{2\omega k} \cos(\omega t - \epsilon) + Ae^{-kt} \cos \{\sqrt{n^2 - k^2}\ t - \alpha\} \qquad (20),$$

where $\tan \epsilon = \dfrac{2\omega k}{n^2 - \omega^2} \quad . \quad . \quad . \quad . \qquad (21)$

The second term represents the free vibration which gradually dies out, leaving permanently the "forced" vibration which is represented by

$$z = \frac{E \sin \epsilon}{2\omega k} \cos(\omega t - \epsilon) \quad . \quad . (22),$$

and which must now be investigated somewhat more closely

If $n > \omega$, *i e* if the forced period be greater than the natural period, ϵ lies in the first quadrant, and the forced vibration is, as regards phase, behind the force If, on the other hand, $n < \omega$, the forced vibration is accelerated as compared with the force

If the forced and free vibrations have the same period, $n = \omega$ and

$$z = \frac{E}{2\omega k} \sin \omega t$$

Here the motion is a quarter of a period behind the force and the amplitude becomes very great for small values of k

If there is very little friction, we may put, neglecting higher powers,

$$\epsilon = \tan \epsilon = \sin \epsilon = \frac{2\omega k}{n^2 - \omega^2},$$

and the equation of motion is

$$z = \frac{E}{n^2 - \omega^2} \cos \omega \left(t - \frac{2k}{n^2 - \omega^2}\right) \qquad (22a)$$

The friction now only affects the phase For vanishing k, the phase is in complete agreement with that of the force when $n > \omega$ and in complete disagreement when $n < \omega$

As a suggestive example of forced vibrations, neglecting friction we may work out the case of one pendulum having mass m, anc length l, suspended from another pendulum of mass M and length L For the equations of motion of m, we have, neglecting friction an confining ourselves to small motions,

$$x + \frac{g}{l}(x - x_1) = 0,$$

where x and x_1 are the displacements of m and M respectively.

If $2\pi/n$ be the free period of m, when M is stationary, the equatio may be written

$$x + n^2 (x - x_1) = 0 \qquad (23)$$

To form the equations of motion of M, we may take the tension the lower string to be mg to the degree of accuracy aimed at Hen writing α for the ratio of the masses, and n_1^2 for the free period of when the lower string is not attached, the equation of motion become

$$x_1 + n_1^2 x_1 + n^2 \alpha (x_1 - x) = 0 \qquad (24)$$

We shall not attempt to obtain a general integral of these equatio but confine ourselves to that particular solution in which each pendulu can perform a simple periodic motion Writing therefore

$$x_1 = a \cos \omega t \qquad (25),$$

$$x = ra \cos (\omega t - \epsilon) \qquad (26),$$

we see at once by substitution that (24) cannot be satisfied unless $\epsilon =$

Substituting (25) and (26) into (23) and (24) we obtain the eq tions of condition

$$-\omega^2 + n_1^2 + n^2 \alpha (1 - r) = 0,$$

$$-r\omega^2 + n^2 (r - 1) = 0,$$

from which r and ω may be obtained.

The second equation gives

$$r = \frac{n^2}{n^2 - \omega^2},$$

and on substitution into the first, we obtain

$$n_1^2 - \omega^2 = \alpha \frac{n^2\omega^2}{n^2 - \omega^2}$$

When M, the mass of the upper pendulum, is great compared with m, α is small and ω then becomes nearly equal to n_1, i.e. the period of the combined pendulum is nearly equal to the period of the upper pendulum, as is indeed to be expected. Substituting $\omega = n_1$ in the terms involving α, we obtain as a second approximation

$$\omega^2 = n_1^2 \left(1 - \alpha \frac{n^2}{n^2 - n_1^2}\right).$$

The combined period is therefore longer than that of the upper pendulum when $n > n_1$, i.e. when the upper pendulum has already a longer period than the shorter one. We must draw therefore the unexpected conclusion that the combined period does not lie between the two free periods, but is greater or less than that of the heavy upper pendulum according as that pendulum has already a greater or less period than the lower one.

If α is small, r approaches the value $n^2/(n^2 - n_1^2)$, and there is agreement or opposition of phase according as n is greater or smaller than n_1, i.e. according as the upper pendulum has the longer or shorter period. The relative positions of the pendulum in the two cases are represented in Fig. 166, the time being such that both pendulums are at their points of greatest deviation.

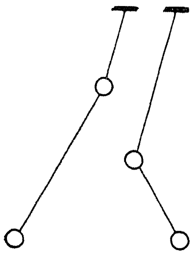

Fig. 166

147. Passage of light through a responsive medium. We now consider light to pass through a medium, the particles of which are subject to forces, capable of giving rise to free vibrations of definite periods. We consider plane waves propagated in the x direction, the displacements being in the z direction. In order to obtain a simply periodic motion for the free vibrations of the particles, we may imagine each to be attracted to a fixed centre by a force varying as the distance. This centre of force we take to form part of the medium to which it is rigidly attached. If ζ be the displacement of the medium, and ζ_1 that of the particle, the equation of motion of the particle is

$$\zeta_1 + n^2 (\zeta_1 - \zeta) = 0 \quad \dots \quad (27).$$

If ρ_1 is the mass of the particle, $n^2\rho_1$ is the force of attraction at unit distance from the centre of force. The reaction of that force has

to be taken into account in forming the equations of motion of the medium At each centre, the medium is acted on by a force $n^2\rho_1(\zeta_1-\zeta)$, and if there are a great many particles within the distance of a wave-length, we may average up the effects and imagine all the forces to be uniformly distributed Let ρ be the inertia of that portion of the medium which contains on the average one and only one particle. Then the equation for the propagation of the wave is

$$\ddot{\zeta}+\beta n^2(\zeta-\zeta_1)-V^2\frac{d^2\zeta}{dz^2}=0 \qquad (28),$$

where we have written $\beta=\rho_1/\rho$ and V stands for the velocity of propagation when there are no particles or when $n=0$.

If the wave is of the simple periodic type

$$\zeta=\cos(ax-\omega t) \qquad (29)$$

and if the motion has continued without disturbance for a sufficiently long time for the free vibrations of the particle to have died out, their position is expressed by

$$\zeta_1=r\cos(ax-\omega t) \qquad (30)$$

x is a parameter which is constant for each particle, but varies from particle to particle By substituting (29) and (30) into (28) and (27), two equations to determine a and r are obtained

$$-\omega^2 r+n^2(r-1)=0,$$

$$-\omega^2-\beta n^2(r-1)+Va^2=0$$

The first equation gives

$$r=\frac{n^2}{n^2-\omega^2} \qquad (31),$$

and the second $$\frac{V^2a^2}{\omega^2}=1+\frac{\beta n^2}{n^2-\omega^2}.$$

ω/a is the velocity (v) of transmission of the wave having a frequency $\omega/2\pi$, so that finally

$$\frac{1}{v^2}=\frac{1}{V^2}\left(1+\frac{\beta n^2}{n^2-\omega^2}\right) \qquad (32)$$

The frequency of the free vibration is $n/2\pi$

This is Sellmeyer's equation, by means of which he first showed that the velocity of light must depend on the periods of free vibration of the molecules embedded in the æther

148. General investigation of the effect of a responsive medium. It will be useful to introduce here a more general investigation, which we shall base on the electromagnetic theory

In Art 134 we had expressed the total current as the sum of a polarization or displacement current and the conduction current. To

this, we may now add the convection currents If there are N positive electrons in unit volume, each carrying a charge e and moving with velocity ζ_1 in the z direction, then $Ne\zeta_1$ is the z component of the convection current, and to this we must add the convection current of negative electricity $-Ne\zeta_2$ We may include both currents in the expression $Ne\zeta$ if ζ denote the relative velocity of positive and negative electricity The conduction current is also due to the convection of electrons, but we leave it in the form CE, because we want to distinguish between the current subject to ohmic resistance which forms a system depending only on one variable, and that which is due to oscillations of electric charges within the molecule. Confining therefore ζ to the velocity of these charges, we have for the z component of the total current in place of (17) Chapter x,

$$w=\frac{1}{4\pi}K\frac{dR}{dt}+CR+Ne\zeta \quad . . \quad (33)$$

The last of equations (12) Chapter x gives with the help of (13) and putting the magnetic permeability equal to one,

$$4\pi\frac{dw}{dt}=\frac{d}{dy}\left(\frac{dR}{dy}-\frac{dQ}{dz}\right)-\frac{d}{dx}\left(\frac{dP}{dz}-\frac{dR}{dx}\right)$$
$$=\nabla^2R-\frac{d}{dz}\left(\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}\right) \quad . . \quad (34)$$

The last term vanishes in isotropic media, and we may in that case, eliminating w, write for the equation of motion in the z direction

$$\frac{d}{dt}\left(K\frac{dR}{dt}+4\pi CR+4\pi Ne\zeta\right)=\nabla^2R \quad (35)$$

Always assuming the disturbance to be simply periodic, the displacement ζ may be divided into two portions, one of which is in phase with R and the other in phase with dR/dt Writing therefore

$$4\pi\zeta=AR-\frac{B}{\omega}\frac{dR}{dt} \quad . \quad (36),$$

and

$$4\pi\dot{\zeta}=A\frac{dR}{dt}-\frac{B}{\omega}\frac{d^2R}{dt^2}$$
$$=A\frac{dR}{dt}+B\omega R \quad . \quad . \quad (37),$$

(35) becomes

$$(K+NAe)\frac{d^2R}{dt^2}+(4\pi C+\omega NBe)\frac{dR}{dt}=\nabla^2R$$

Comparing with this (1), we see that the investigation of Art 144 applies to this case, writing

$$G=K+NAe,$$
$$F=4\pi C+\omega NBe.$$

Hence by (13) $$\nu^2 - \kappa^2 = (K + NAe)\, V^2 \qquad (38),$$

$$2\nu\kappa \cos\rho = (2C\lambda + NBVe)\, V \qquad (39)$$

If the specific inductive capacity of the intermolecular space is the same as that of empty space, $KV^2 = 1$, and (38) is replaced by

$$\nu^2 - \kappa^2 = 1 + NAeV^2 \qquad (40)$$

149. Wave velocity in a responsive medium according to the electromagnetic theory. P Drude was the first to apply the electromagnetic theory to the explanation of dispersion based on the principles of sympathetic vibrations It is now generally accepted that each molecule contains a number of electrons, and that each electron consists of a definite electric charge concentrated within a space which must be small compared with the size of the molecule Every quantity of electricity is made up of these electrons just as any ordinary substance is made up of molecules It is therefore correct to speak of an electron as an "atom" of electricity A moving electron represents kinetic energy which exists in the medium surrounding the electron in virtue of the magnetic field established by the motion This energy is proportional to the square of the velocity, and may be expressed therefore in the form $\frac{1}{2}\rho v^2$ if v be the velocity of the electron The quantity ρ we may call the apparent inertia of the electron Students must however guard against being misled into the error of believing that if there are several electrons near each other, their total kinetic energy may be written down as a sum of their separate kinetic energies That total energy contains products of the form $v_1 v_2$, where v_2 and v_1 refer to distinct electrons While dealing with a single electron we are, however, justified in applying to it the ordinary laws of dynamics substituting ρ for its mass If we therefore consider that the incident vibrations of light excite the sympathetic motion of a single electron in a molecule, we may write for the equation of motion

$$\zeta + n^2\zeta - \frac{eR}{\rho} = 0$$

In forming this equation, we have imagined the electron to be acted on by a centre of force varying as the distance, the force being $n^2\rho$ at unit distance, while e is the charge of the electron and R the external electric force acting on it If R varies as $e^{-\iota\omega t}$ and the free vibration has died out, so that the period of ζ is that of the incident force, we may deduce

$$\zeta = \frac{eR}{\rho\,(n^2 - \omega^2)}$$

By comparison with (36) we see that in the present case $B = 0$, and

$$A = \frac{4\pi e}{\rho\,(n^2 - \omega^2)} \qquad (41)$$

If the medium is a non-conductor, (39) shows that $\kappa = 0$ and ν in that case is equal to the refractive index μ. We therefore finally obtain from (40)

$$\mu^2 = 1 + \frac{4\pi e^2 N V^2}{\rho (n^2 - \omega^2)} \qquad (42)$$

If the charge e is measured in the electrostatic system, we may leave out the factor V^2 in the second term of the right-hand side, and the equation is then identical with that given by Drude*. The dimensions of the equation are easily checked, as e^2/ρ is of the dimension of a length.

The way in which the refractive index changes with a change in the wave-length of the incident light, is easily recognized by inspection of (42). For long waves, ω is small, and in the limit when $\omega = 0$,

$$\mu = 1 + \frac{4\pi e N V^2}{\rho n^2}$$

With increasing ω the refractive index increases until $n = \omega$. At that point there is a discontinuity, μ^2 suddenly changing from $+\infty$ to $-\infty$.

For a definite value of ω larger than n, μ is zero and afterwards μ increases once more and approaches unit value for infinitely short periods. If we introduce K' the specific inductive capacity of the medium which must coincide with the value of μ^2 for $\omega = 0$, we may write

$$\mu^2 = K' + \frac{4\pi e^2 N V^2 \omega^2}{\rho (n^2 - \omega^2)} \qquad (43).$$

Fig 167 represents the curve $y = \dfrac{1}{1 - x^2}$. By a change in the vertical scale, and a displacement of the horizontal axis, the curve may be made to coincide with that which gives the relation between μ^2 and ω. In order to establish agreement with the subsequent curves in which small values of ω are placed on the right, the positive axis of x in the figure is drawn towards the left.

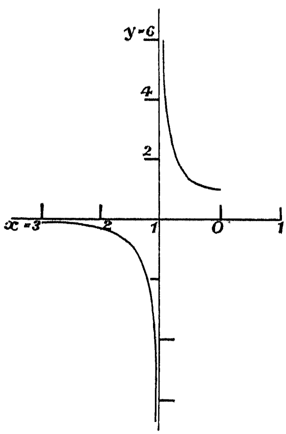

Fig 167

The above investigation gives the essential features of the theory of dispersion in its simplest form, and may be extended so as to approach more nearly the actual conditions. Within each molecule there are a number of free periods of vibration as is shown by spectroscopic

* *Lehrbuch der Optik*, Chapter v.

observation. It is necessary to consider therefore a connected system of electrons, the state of motion for each free period being defined by one variable Assuming these periods to be independent of each other, let ψ_1, ψ_2, etc be the variables For small changes $\delta\psi_1$, $\delta\psi_2$, of these variables, the total work done may be expressed in the form

$$\Psi_1\delta\psi_1 + \Psi_1\delta\psi_2 + \quad = 0$$

If the displacements from the stable configuration are small, an equation of motion holds for each variable of the form

$$\psi_m + n_m^2\psi_m = \Psi_m^{*},$$

where Ψ_m is the generalized component of force. The frequency of the free period being $n_m/2\pi$, and the motion being assumed to be proportional to $e^{-\iota\omega t}$, the equation may be written

$$(n_m^2 - \omega^2)\psi_m = \Psi_m$$

If the external electric force is $Re^{-\iota\omega t}$, we have for the different variables

$$(n_1^2 - \omega^2)\psi_1 = \Psi_1 = A_1R,$$
$$(n_2^2 - \omega^2)\psi_2 = \Psi_2 = A_2R,$$

etc where A_1 and A_2 are constants Linear relationships must hold between the displacements ζ_1, ζ_2, etc of the electrons and ψ_1, ψ_2, so that for the total current within each molecule, we have

$$e(\zeta_1 + \zeta_2 + \zeta_3 + \quad) = e(a_1\psi_1 + a_2\psi_2 + \quad)$$
$$= e\frac{dR}{dt}\left(\frac{a_1A_1}{n_1^2 - \omega^2} + \frac{a_2A_2}{n_2^2 - \omega^2} + \quad\right) \quad . \qquad (44)$$

Hence A in (37) becomes

$$A = \Sigma\frac{a_mA_m}{n_m^2 - \omega^2} \qquad \ldots \quad . \ (45)$$

a_1, a_2, are a set of constants, the connexion between which and A_1, A_2 etc it is not necessary to discuss here

From the equation defining the electric current, we may, as in the simple case which has been treated in detail, derive the expression corresponding to (42) which now takes the form

$$\mu^2 = 1 + \frac{\beta_1}{n_1^2 - \omega^2} + \frac{\beta_2}{n_2^2 - \omega^2} + \qquad . \qquad (46),$$

where β_1, β_2, etc are numerical constants defining the dispersion of the medium For waves of infinitely long periods, we have

$$\mu_\infty^2 = K' = 1 + \frac{\beta_1}{n_1^2} + \frac{\beta_2}{n_2^2} + \frac{\beta_3}{n_3^2} + . \qquad . \quad . \qquad (47).$$

If there is only one period of vibration, β_1 may be determined from the observed inductive capacity, and if K' is known, the refractive index for all waves is completely determined

* Lord Rayleigh, *Theory of Sound*, Vol I Chapter IV

By using (47) equation (46) may be written, changing the constants,

$$\mu^2 = K' + \Sigma \frac{M_m}{\lambda^2 - \lambda_m^2} \qquad \qquad (48),$$

where λ is the wave-length in vacuo to which ω applies, and λ_m is the wave-length, also measured in vacuo, which corresponds to the free vibration of the molecules

150 Dispersion in transparent media. If the range of spectrum considered is far removed from any of the free periods of the molecule, the dispersion formula may conveniently be expressed differently If the region of resonance lies in the ultra-violet, we may expand in terms of a series proceeding by $\frac{\lambda_m}{\lambda}$ and thus find

$$\mu = K' + \frac{A_1}{\lambda^2} + \frac{A_2}{\lambda^4} + \qquad \qquad (49),$$

where $\quad A_1 = \Sigma M_m, \; A_2 = \Sigma M_m \lambda_m^2, \; A_p = \Sigma M_m \lambda_m^{2p-2}.$

Equation (49) is known by the name of Cauchy's formula, but was deduced by Cauchy in quite a different manner

If the region of resonance is in the infra-red, we may express the series in terms which proceed by $\frac{\lambda}{\lambda_m}$ and thus obtain

$$\mu^2 = K' - B_0 - (B_1\lambda^2 + B_2\lambda^4 + \dots) \qquad \qquad (50),$$

where $\quad B_0 = \Sigma \frac{M_m}{\lambda_m^2}, \; B_1 = \Sigma \frac{M_m}{\lambda_m^4}, \; B_2 = \Sigma \frac{M_m}{\lambda_m^6}, \; B_n = \Sigma \frac{M_m}{\lambda_m^{2p+2}}.$

In the case of many substances Cauchy's formula does not give a sufficient representation of the actual dispersion without the addition of a negative term proportional to λ^2. This fact which has been clearly established by Ketteler suggests that though the dispersion in the visible part is mainly regulated by ultra-violet resonance, it is also to some extent influenced by free periods lying in the infra-red. Assuming for the sake of simplicity one infra-red and one ultra-violet free period, having wave-lengths λ_r and λ_v respectively, equation (48) becomes

$$\mu^2 = K' + \frac{M_v}{\lambda - \lambda_v^2} + \frac{M_r}{\lambda_r^2 - \lambda^2} \qquad \qquad (51).$$

This equation has been tested over a long range of wave-lengths for rock-salt, sylvin and fluorspar, and the agreement arrived at is sufficient to show that in its essential points, the present theory is correct, and that refraction is a consequence of the forced vibration of the molecules, which respond strongly to the periodic impulses of those waves which are in sympathy with its periods of free vibration These experiments, which are fundamental in the theory

of refraction, have been made possible by the beautiful device of H Rubens and E F Nichols*, which enabled them to obtain fairly homogeneous radiations of large wave-lengths by multiple reflexion The success of the method itself is an excellent confirmation of the above theory

As will appear in the next article, a substance will totally reflect the radiations having periods equal to that of the free vibrations of the molecules By successive reflexion from a number of surfaces, all wave-lengths are eliminated except those for which there is approximately total reflexion. It was found in this manner that with quartz, wave-lengths of 20 75 and 8 25 mikroms, and with fluorspar a wave-length of 23 7, could be obtained

The refractive index of quartz is represented with considerable accuracy by the formula

$$\mu_2 = K' + \frac{M_v}{\lambda^2 - \lambda_v^2} - \frac{M_r}{\lambda_r^2 - \lambda^2} - \frac{M_s}{\lambda_s^2 - \lambda^2},$$

in which λ_r and λ_s are directly determined by observation

Rubens and Nichols also determined the refractive indices of rock-salt and sylvin for the wave-lengths 20 75μ and 8 85μ and hence could deduce an equation to represent the dispersion of these two substances through a wide range In the following Table, I have collected the constants of the substances used by the authors, adding Paschen's† numbers for fluorspar

TABLE X

	Quartz	Fluorspar	Rock-salt	Sylvin
M_v	01065	00612	01850	0150
M_r	44 224	5099	8977	10747
M_s	713 55	—	—	—
λ_v	1031	09425	127	153
λ_r	8 85	35 47	56 12	67 21
λ_s	20 75			
K'	4 5788	6 0910	5 179	4 553
K	4 55	6 8	5 85	4 94

All wave-lengths are given in mikroms, *i e* in 10^{-4} cms As has been stated, the resonance periods of quartz have been derived from observation, the others are calculated from the dispersion

* *Wied Ann* LX p 418 (1897).
† *Wied Ann* LIII p 812

formula A good confirmation was subsequently obtained by Rubens and Aschkinass* in the direct determination of the resonance region in rock-salt and sylvin, though the observed free periods were not found to coincide as much as might have been wished with those derived from the dispersion formula The wave-lengths for total reflexion were measured to be 51 2 and 61 1 instead of 56 1 and 67 2 as given in the table There is still less agreement in the case of fluorspar, the wave-length best fitting the observation being 35 47, while the region of total reflexion lies at 23 7 The discrepancy may be due to the fact that as in quartz, fluorspar has a second region of total reflexion in the infra-red

The next remark called for by the inspection of the table is connected with the relative small values of the constants M in the ultra-violet term This must be due to the comparative smallness of the resonance for short periods There is a gradual increase of the value of M for diminishing values of the resonance period This increase is not very uniform, but is such that in general it is more rapid than the increase in the square of the wave-length at which resonance takes place A closer investigation of this point seems called for, but it would be necessary for the purpose to take account of the molecular volumes of the different substances If the refractive index for infinitely short waves is one, as required by (46), equation (48) shows that the constants should satisfy the condition

$$\Sigma \frac{M_m}{\lambda_m^2} = K' - 1$$

This relation is only approximately fulfilled by the numbers given in the Table, but its complete verification was not to be expected considering that there are probably unknown regions of resonance in the ultra-violet

The constant K' should, according to theory, be equal to the specific inductive capacity of the medium, the last two rows of the table show that though the present agreement is not by any means perfect, its power to represent the facts is a considerable stage in advance of the older theories which gave Cauchy's formula (Compare Art 136)

151. Extension of the theory. Our theory requires extension in two directions Sellmeyer's equation

$$\mu^2 - 1 = \Sigma \frac{\beta_m n_m^2}{n_m^2 - \omega^2}$$

gives infinitely large values of μ whenever the period of the incident

* *Wied Ann* LXV p 241 (1898)

light coincides with one of the free periods of vibration This is a consequence of the infinite amplitude of the forced vibration, as it appears e g in equation (22a)

These infinite amplitudes may be avoided by the introduction of a frictional term retarding the free vibrations as in the case treated in Art 144 Real friction is not admissible in the treatment of molecular vibrations, but as there is loss of energy due to radiation, there must be some retarding force, which is in phase with the velocity Its effect will therefore be the same as that of a frictional force To find the constants determining the wave velocity in this case we apply equation (20) Confining ourselves for simplicity to a single variable, we may write for the displacement of the electrons in the molecule

$$\zeta = \frac{R_0 e \sin \epsilon}{2\rho\omega k} \cos(\omega t - \epsilon),$$

where $R = R_0 \cos \omega t$ represents the electric force due to the incident light and ρ has the same meaning as in Art 149

Introducing R in place of R_0 the equation becomes

$$\zeta = \frac{e \sin 2\epsilon}{4\rho\omega k} R - \frac{e \sin^2 \epsilon}{2\rho\omega^2 k} \frac{dR}{dt},$$

where by (21) $\tan \epsilon = \frac{2\omega k}{n^2 - \omega^2}$

Hence by comparison with (36)

$$A = \frac{\pi e \sin 2\epsilon}{\rho\omega k} = \frac{2\pi e}{\rho\omega k} \frac{\tan \epsilon}{1 + \tan^2 \epsilon}$$

$$= \frac{4\pi e}{\rho} \frac{n^2 - \omega^2}{4\omega^2 k^2 + (n^2 - \omega^2)^2} \quad . \quad (52),$$

and

$$B = \frac{2\pi e \sin^2 \epsilon}{\rho\omega k} = \frac{2\pi e}{\rho} \frac{\tan^2 \epsilon}{1 + \tan^2 \epsilon}$$

$$= \frac{2\pi e}{\rho} \frac{4\omega k}{4\omega^2 k^2 + (n^2 - \omega^2)^2}$$

Hence (39) and (40) become, taking account of (12), and putting $C = 0$,

$$\left.\begin{aligned} \nu_0^2 - \kappa_0^2 &= 1 + 4\pi N e^2 V^2 \frac{n^2 - \omega^2}{\rho\{4\omega^2 k^2 + (n^2 - \omega^2)^2\}} \\ \nu_0 \kappa_0 &= \quad 2\pi N e^2 V^2 \frac{2\omega k}{\rho\{4\omega^2 k^2 + (n^2 - \omega^2)^2\}} \end{aligned}\right\} \quad . \quad (53).$$

If κ_0 be small, so that κ_0^2 may be neglected, ν_0 becomes equal to the refractive index of the substance, which then refracts according to the sine law The introduction of k has got rid of the infinite value of ν_0, but the value of k will be shown in Chapter XIII to be too small to be the cause of the observed absorption phenomena

152. Finite range of Free Vibrations. In the case of solids and liquids, we may judge by their absorption effects that the free vibrations are not confined to definite wave-lengths, but extend over a finite range. In order to extend the theory so as to include this case we neglect k and we write βdn in (42) for $4\pi e^2 NV^2/\rho$. We then find the total effect of the forced vibration by integrating over the absorption range, thus

$$\mu^2 - 1 = \int_{n_1}^{n_2} \frac{\beta dn}{n^2 - \omega^2}.$$

Here β may be a function of n. Assuming it to be constant, we find on integration

$$\mu^2 - 1 = \frac{\beta}{2\omega} \log \left| \frac{(n_2 - \omega)(n_1 + \omega)}{(n_2 + \omega)(n_1 - \omega)} \right| \quad . \quad . \quad (54),$$

where the absolute value of the fraction, the logarithm of which occurs in the expression, is to be taken. The square of the refractive index is infinitely large on the positive side for $\omega = n_1$, and infinitely large on the negative side for $\omega = n_2$, n_2 belonging to the higher frequency. The region of ω for which μ^2 is negative includes that range of waves which cannot enter the medium. It is bounded on the red side by the value of ω for which

$$\log \frac{(\omega - n_1)(n_2 + \omega)}{(n_2 - \omega)(n_1 + \omega)} = \frac{2\omega}{\beta},$$

and on the side of higher frequency by the value of ω for which

$$\log \frac{(n_1 - \omega)(n_2 + \omega)}{(n_1 + \omega)(n_2 - \omega)} = \frac{2\omega}{\beta}.$$

The infinity of the refractive index at the lower frequency edge of the absorption is avoided if the intensity of the absorption band is assumed to diminish gradually to zero on both sides instead of beginning and ending abruptly. As a simple example we may take the case that the absorption between n_1 and n_2 is equal to $\beta(n - n_1)(n_2 - n)$. The expression for $\mu^2 - 1$ then becomes

$$\mu^2 - 1 = \int_{n_1}^{n_2} \frac{\beta(n - n_1)(n_2 - n)}{n^2 - \omega^2} dn$$

$$= \beta \int_{n_1}^{n_2} \left\{ \frac{(n_1 + \omega)(n_2 + \omega)}{2\omega(n + \omega)} - \frac{(n_1 - \omega)(n_2 - \omega)}{2\omega(n - \omega)} - 1 \right\} dn$$

$$= \beta \left\{ \frac{(n_1 + \omega)(n_2 + \omega)}{2\omega} \log \frac{n_2 + \omega}{n_1 + \omega} \right.$$

$$\left. - \frac{(n_1 - \omega)(n_2 - \omega)}{2\omega} \log \left| \frac{n_2 - \omega}{n_1 - \omega} \right| + n_1 - n_2 \right\} \quad \ldots \quad .. \; (55).$$

In the second term the sign has to be chosen so as to give always

a positive value to the fraction At the edges of the band, we have for $\omega = n_1$

$$\mu^2 - 1 = \beta \left\{ (n_1 + n_2) \log \frac{n_1 + n_2}{2n_1} - (n_2 - n_1) \right\},$$

for $\omega = n_2$

$$\mu^2 - 1 = \beta \left\{ (n_1 + n_2) \log \frac{2n_2}{n_1 + n_2} - (n_2 - n_1) \right\}$$

153 Absorption Absorption has so far not been taken into account though in Art (151) the coefficient k involves a gradual extinction of light as the wave proceeds This extinction was made to depend on the imparting of energy from the vibrating electron to the medium The particle takes up energy from the incident light and communicates it to the æther in the form of a vibration having the same period as the incident light This is a case of scattering of light and not of absorption Analytically, of course, we may include absorption if the value of k is increased beyond the amount required for the scattering This has been very generally done by various writers, but there is no theoretical justification for it.

Absorption is not brought about by anything analogous to ordinary friction, but most probably by sudden changes in the vibratory motion due to molecular impacts We may take as typical, equation (20) Art 146, which represents the motion of a vibrating particle subject to a periodic force The second term quickly subsides if the particle is left to itself, but is constantly being renewed by molecular impacts. The free vibration therefore never disappears and its energy is derived from that of the incident vibration It is worth while looking a little more closely at this question Absorption means the gradual decay of the forced vibration into the vibrations and translatory motions corresponding to the temperature of the body, and this takes place in the first instance by a change of the forced vibration into the vibrations of the free periods Molecular impacts sufficiently account for this first step Two points should be specially noted The forced vibration as indicated by the first term in (20) is not disturbed by the impacts Its phase and amplitude remain the same and persist throughout the motion Secondly,—and this is very important—*the free vibration does not represent a simple periodic motion* The waves sent out are the less homogeneous the more quickly the free vibrations decay Molecular impacts therefore increase the possible free periods of motion. If luminous the body would radiate in periods extending over a finite range on both sides of $2\pi/\sqrt{n^2 - k^2}$, and by the principle of exchange, it must be capable of absorbing those periods which it can radiate.

The periods contained in the second term of (20) include $2\pi/\omega$ which is that of the force, and this has to be taken into account in judging of the effect of the free periods on absorption and wave velocity From the principle of conservation of energy, we may predict that the effect of that portion of the free vibration which has a period $2\pi/\omega$ will be to reduce the amplitude of the forced vibration, because it is the energy of the forced vibration which by the supposed impact supplies the energy of the rest of the motion. Hence instead of A, as given in (41), we must introduce a value reduced in some unknown ratio. This does not affect our final equations (48) as they contain undetermined factors M_1, M_2

Again the principle of conservative energy tells us that the transmitted light must be diminished by molecular impacts, hence the second term of (20) must contain a vibration of period $2\pi/\omega$ which as regards phase is a right angle behind that of the incident light While B in the simple theory of Art 149 was zero, we must now conclude that one effect of molecular impact is the introduction of a real positive value for B That value need not however be that arrived at in Art 151 by the assumption of a frictional term.

For non-conductors, for which $C=0$, equations (39) and (40) become

$$\left.\begin{aligned} 1+NAeV^2 &= \nu_0^2-\kappa_0^2 \\ NBeV^2 &= 2\nu_0\kappa_0 \end{aligned}\right\} \qquad (56).$$

If κ_0 is determined by observation, the first of these equations gives

$$\nu_0^2-1=\kappa_0^2+NAeV^2,$$

or substituting the more general value of A as determined by (45),

$$\nu_0^2-1=\kappa_0^2+\Sigma\frac{\beta_m}{n_m^2-\omega^2} \qquad (57).$$

If the absorption band ranges over a finite width between values of $n=n_1$ and $n=n_2$, the absorption being proportional to $(n-n_1)(n-n_2)$, we may substitute the right-hand side of (54) for the second term of (57). If κ_0 is to be obtained from the general theory, both equations (56) must be used. Eliminating κ between them, we find for the general equation of ν_0

$$\nu_0^2-\frac{N^2B^2e^2V^4}{4\nu_0^2}=1+NAeV^2 \qquad (58).$$

If we substitute terms of the form (52) instead of (45) for A, we take account of the loss of energy by radiation, and have then obtained the most general theory of refraction that can at present be formed. The quantity B is not known, but we shall probably not go far wrong in taking it to be proportional to A.

154. Selective refraction. The phenomena which have called forth the theoretical discussions of this Chapter have been grouped together under the name "Anomalous Dispersion" But we are now prepared to say that there is nothing anomalous in the effect of absorption on refraction, and that the ordinary or "normal" dispersion is only a particular case of the "anomalous" one Under these circumstances the name is misleading, and I therefore introduce the more appropriate one of "Selective Refraction" and "Selective Dispersion"

The experimental illustration of selective refraction is rendered somewhat difficult by the fact that the substances which show the effects are all highly absorbent With a hollow prism filled with a strong solution of fuchsin or cyanin, it may easily be demonstrated that the red of the spectrum is more refracted than the violet, but dispersion in the immediate neighbourhood of the absorption band is too great to make exact measurements in that region possible. Kundt originated a method of observation which is often employed The vertical slit of a spectroscope is illuminated by projecting upon it the image of a *horizontal* slit, through which white light is passed If the horizontal slit be narrow, an almost linear spectrum is seen, running along a horizontal line The position of this horizontal line may be marked If now a prism filled with a substance showing selective refraction be interposed between the horizontal slit and its image, the refracting edge of this prism being horizontal and downwards, the line of the observed spectrum will no longer be straight Were the prism filled with water, the spectrum would run upwards in a curved line from red to violet A curve running downwards from red to violet would indicate a refractive index diminishing with increasing frequencies Refractive indices smaller than one, showing a velocity of light greater than that of empty space, would be indicated by displacement below that of the original linear spectrum For purposes of illustration, and for measurements when the angle of the prism is small, this method is very successful.

The simplest case of selective refraction is shown by sodium vapour, as the absorption is here confined to two regions, each of which covers only a very narrow range of wave-lengths In other words, the refraction is affected by absorption "lines" as distinguished from absorption bands The selective refraction of a luminous conical sodium flame was first shown by Kundt, who however did not investigate the specially interesting region which lies between the absorption lines This has been done by Becquerel Plate II Fig 6 is a reproduction of one of Mr Becquerel's photographs, the red end being to the right The sodium vapour was used in the form of a luminous flame formed by a special device into a prismatic shape

The horizontal black line marks the horizontal linear spectrum in its original position. The horizontal portion of the white band, the centre of which is slightly raised above the black line, shows that at a short distance from the double sodium line there is a slight displacement upwards indicating a refractive index somewhat greater than one. The nearly vertical branches of the curve indicate a considerable dispersion close to and between the absorption lines. The course of this dispersion is better studied in Fig. 168, which has been drawn from the measurements given by Mr Henri Becquerel*, upward displacements being approximately proportional to $\mu-1$. It will be seen that in accordance with the previous theory, the refractive index rapidly increases as we approach each absorption line from the red end, and that the light which vibrates just a little more quickly than that corresponding to the absorption band has its velocity increased. Mr Becquerel calculates approximately that the light in close proximity to D_2 and on its violet side has a refractive index of ·9988, so that its velocity is about ·1 % greater than that in empty space. Concentrated solutions of colouring matters exhibit the phenomena of selective refraction, but here the theory is complicated by the fact that the absorption extends over a wide range of wave-lengths. Some of these colouring matters may be shaped into solid prisms of small angle, by means of which the refractive indices for different periods and the coefficients of absorption may be measured. Pflüger† takes a few drops of a concentrated solution of the colouring matters in alcohol, and runs the solution into the two wedge-shaped spaces between a glass plate and a wide glass tube. As the solvent evaporates, it leaves behind a double prism. Amongst many prisms made in this fashion, a few may be found with surfaces sufficiently good to render optical investigation of refractive indices possible. The prisms used by Pflüger had a refracting angle of from 40—130 seconds of arc, and the refractive indices could be determined throughout the absorption band. It is a special merit of Pflüger's

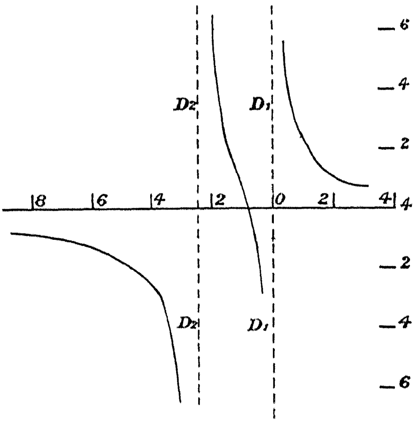

Fig. 168

* *Comptes Rendus* CXXVIII. p. 145 (1899).

† *Wied. Ann.* LVI. p. 412 (1895) and LXV. p. 173 (1898).

investigations that he determined also the coefficients of absorption for the different wave-lengths As a thickness of very few wave-lengths is sufficient to extinguish the light, the plates used for the purpose had a thickness of less than half a wave-length In Figure 169

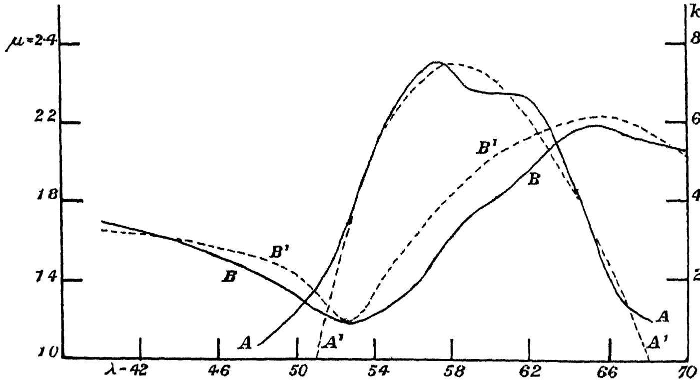

Fig 169

Pflugér's curves for cyanin are reproduced, the curves of refractive index and coefficient of extinction being marked B and A respectively I have added in dotted lines an assumed absorption curve following the law discussed in Art 150, the constants being roughly adjusted so as to fit the edges of the absorption band A second dotted line shows the curve of refraction (B') calculated from equation (55) after substitution of $\nu^2 - \kappa^2$ for μ^2. The value of β was determined so as to give roughly the proper quantity for the difference in the refractive indices near the two edges of the absorption band, and a constant term has been added to represent the effect of infra-red and ultra-violet absorption The correspondence between calculated and experimental values might be made closer if instead of a constant term, one varying with the wave-length had been taken, but in view of the fact that the assumed law of absorption only approximately represents the facts, it seems unnecessary to seek for a closer agreement of the refraction curves The more sudden fall and rise of the calculated dispersion curve near the green boundary of the absorption band is due to the fact that the actual absorption curve does not show the rapid increase of absorption indicated by the assumed curve

155. Metallic Reflexion We include under the term "metallic" reflexion, all cases in which the greater portion of the incident light is returned, in consequence, as it will appear, of the absorptive power of the medium If the amplitudes of the incident, reflected and

refracted light be denoted by 1, r, s, respectively, we may as in Art 140 write

for the incident wave
$$e^{\iota(ax+by-\omega t)},$$
for the reflected wave
$$re^{\iota(-ax+by-\omega t)} \quad\ldots\ldots\ldots(59),$$
and for the disturbance entering the medium
$$se^{\iota(a_1x+by-\omega t)} \quad\ldots\ldots\ldots(60).$$

The surface conditions are the same as for transparent media, hence the previous investigations apply as far as the analytical expressions of r and s are concerned. When the incident light is polarized in the plane of incidence, the amplitude r_s was found to be

$$r_s = \frac{\sin(\theta_1-\theta)}{\sin(\theta_1+\theta)} \quad\ldots\ldots\ldots(61),$$

and for light polarized perpendicularly to the plane of incidence

$$r_p = \frac{\tan(\theta_1-\theta)}{\tan(\theta_1+\theta)} \quad\ldots\ldots\ldots(62)$$

But θ_1 being now complex, r_s and r_p are complex quantities from which we may separately deduce the real amplitude and the change of phase Writing $r_s = h_s e^{\iota\delta}$ and substituting in (59) we see that h_s is the real amplitude, and δ denotes the change of phase.

It therefore becomes a problem of algebraic transformation to change (61) and (62) into the standard form $he^{\iota\delta}$ Little experimental work has been done to measure the intensity of the reflected light and we may therefore confine ourselves to calculating the value of h for perpendicular incidence for which, as for transparent media,

$$r_s = -r_p = \frac{\mu-1}{\mu+1}.$$

But from (16)

$$\mu = \nu_0 + \iota\kappa_0,$$

$$r_s = he^{\iota\delta} = \frac{\nu_0 - 1 + \iota\kappa_0}{\nu_0 + 1 + \iota\kappa_0} \quad\ldots\ldots\ldots(63)$$

If this quantity be called P, and Q be that obtained from P by reversing the sign of ι, the proposition proved in Art 8 shows that

$$h^2 = PQ, \quad \tan\delta = \frac{P-Q}{\iota(P+Q)}.$$

Hence
$$h^2 = \frac{(\nu_0-1)^2+\kappa_0^2}{(\nu_0+1)^2+\kappa_0^2}, \quad \tan\delta = \frac{2\kappa_0}{\nu_0^2+\kappa_0^2-1} \quad\ldots\ldots\ldots(64).$$

It will be noticed that great absorbing power means a large intensity of reflected light, for if κ_0 is large compared with ν_0+1, h is nearly 1, and the light is almost totally reflected. The absorbing power therefore is

active not only in transforming the energy that enters, but also in preventing the light from entering. There is also nearly total reflexion when ν_0 is either large or small compared with 1, but this effect is not confined to opaque substances

When light polarized at an angle ϕ to the plane of incidence falls on a metal surface, it may be considered to be made up of light having amplitude $\cos\phi$ polarized in that plane, and light of amplitude $\sin\phi$ polarized at right angles.

We write R_p and R_s for the reflected amplitudes, which are respectively equal to $r_p \cos\phi$ and $r_s \sin\phi$, where r_s and r_p are defined by (61) and (62)

Putting $$R_s = h_s \cos\phi e^{i\delta_1}, \quad R_p = h_p \sin\phi e^{i\delta_2}$$

and writing $$\tan\psi = h_p/h_s,$$

$$\frac{R_p}{R_s} = \tan\phi \tan\psi \, e^{i(\delta_2 - \delta_1)},$$

$\delta_2 - \delta_1$, for which we shall write δ, represents the difference in phase of the two components of reflected light That light is therefore polarized elliptically If by some compensating arrangement the difference of phase be destroyed, and the reflected light restored to plane polarization, the angle χ which the plane of polarization of the reflected light forms with the plane of incidence, is determined by

$$\tan\chi = \tan\phi \tan\psi$$

The quantities χ and ϕ may be measured, and hence ψ determined. We must endeavour now to express the optical constants ν_0 and κ_0 in terms of δ and ψ For this purpose, we have from (61) and (60),

$$\frac{\cos(\theta_1 + \theta)}{\cos(\theta_1 - \theta)} = \frac{r_p}{r_s} = \frac{h_p e^{i\delta_2}}{h_s e^{i\delta_1}}$$

$$= \tan\psi e^{i\delta}$$

Also

$$\frac{\cos(\theta_1 + \theta)}{\cos(\theta_1 - \theta)} = \frac{\cos\theta_1 \cos\theta - \sin\theta_1 \sin\theta}{\cos\theta_1 \cos\theta + \sin\theta_1 \sin\theta} = \frac{\cot\theta_1 \sin\theta - \tan\theta \sin\theta}{\cot\theta_1 \sin\theta + \tan\theta \sin\theta}$$

Hence introducing (16)

$$\tan\psi e^{i\delta} = \frac{m - \tan\theta \sin\theta + i\kappa}{m + \tan\theta \sin\theta + i\kappa} \qquad (65).$$

This expression is of the same form as (63), and by subjecting it to the same transformation, we find

$$\tan^2\psi = \frac{(m - \tan\theta \sin\theta)^2 + \kappa^2}{(m + \tan\theta \sin\theta)^2 + \kappa^2}$$

Applying $$\cos 2\psi = \frac{1-\tan^2\psi}{1+\tan^2\psi},$$
this reduces to

$$\cos 2\psi = \frac{2m\cos\theta\sin^2\theta}{(m^2+\kappa^2)\cos^2\theta+\sin^4\theta} \quad \text{. . . (66)}$$

Also $$\tan\delta = \frac{2\kappa}{\nu^2+\kappa^2-\tan^2\theta\sin^2\theta} \quad \text{. . . . (67)}$$

If the optical constants κ_0 and ν_0 are known we may use equations (13) and (17) to calculate ν, κ and m, and hence $\cos 2\psi$ and $\tan\delta$ may be found.

When the difference in phase, δ, is equal to a right angle, $\tan\delta$ is infinitely large, and hence in that case

$$\sin^2\theta\tan^2\theta = \nu^2+\kappa^2.$$

This particular value of θ is called the principal angle of incidence, and, for $\kappa = 0$, corresponds to the polarizing angle

The problem as it generally presents itself, consists in determining the optical constants of the metal from observation of ψ and δ; for this purpose (66) and (67) are not convenient, and we must transform (65) in a different manner

We easily deduce from that equation

$$\frac{m+i\kappa}{\sin\theta\tan\theta} = \frac{1+\tan\psi e^{i\delta}}{1-\tan\psi e^{i\delta}},$$
$$= \frac{1+\tan\psi\cos\delta+i\tan\psi\sin\delta}{1-\tan\psi\cos\delta-i\tan\psi\sin\delta},$$

and the right side of this equation may be transformed in the same manner as (63). We thus find

$$m = \frac{\sin\theta\tan\theta\cos 2\psi}{1-\sin 2\psi\cos\delta} \quad \text{. . . . . (68)},$$

$$\kappa = \frac{\sin\theta\tan\theta\sin 2\psi\sin\delta}{1-\sin 2\psi\cos\delta} \quad \text{. . . . . . . (69)}.$$

If instead of using a compensator similar to Babinet's, the elliptic path of the disturbance of the reflected ray is analysed by a quarter wave-plate or Fresnel's rhomb, the quantities measured are the ratio of the axes of the ellipse and the inclination of these axes to the plane of incidence Calling $\tan\Psi$ the ratio of the minor to the major axis, and γ the angle between the major axis and the plane of incidence, we obtain with the assistance of (18) (19) and (20) of Chapter I.,

$$\left.\begin{aligned} m &= \frac{\sin\theta\tan\theta\cos 2\Psi\cos 2\gamma}{1-\cos 2\Psi\sin 2\gamma} \\ \kappa &= \frac{\sin\theta\tan\theta\sin 2\Psi}{1-\cos 2\Psi\sin 2\gamma} \end{aligned}\right\} \quad \text{. . . . . . . (70)}.$$

Having obtained m and κ we determine the optical constants ν_0 and κ_0 in the following manner Equations (12) may be written

$$m^2 - \kappa^2 + \sin^2\theta = \nu_0^2 - \kappa_0^2,$$

$$m\kappa = \nu_0\kappa_0$$

Solving these equations for ν_0 and κ_0 we obtain

$$\left.\begin{aligned} 2\nu_0^2 &= \sqrt{(m^2-\kappa^2+\sin^2\theta)^2+4m^2\kappa^2}+(m^2-\kappa^2+\sin^2\theta) \\ 2\kappa_0^2 &= \sqrt{(m^2-\kappa^2+\sin^2\theta)^2+4m^2\kappa^2}-(m^2-\kappa^2+\sin^2\theta) \end{aligned}\right\} \quad (71).$$

Equations (68), (69) and (71) constitute the solution of our problem in the form in which Ketteler* first gave it This form is to be preferred to the earlier one given by Cauchy, whose solution did not directly lead to the separation of the constants ν_0 and κ_0 but only to a set of equations which involved intermediate constants and variables, having no physical meaning.

In the case of metals, $m^2 + \kappa^2$ exceeds $\sin^2\theta$ sufficiently to allow us generally to neglect the square of $\sin^2\theta/(m^2+\kappa^2)$ Under these circumstances, expressing the square root which occurs in equations (71) as a series proceeding by powers of $\sin^2\theta$, and neglecting all powers higher than the first, we find

$$\left.\begin{aligned} \nu_0 &= m\left(1+\frac{\sin^2\theta}{2(m^2+\kappa^2)}\right) \\ \kappa_0 &= \kappa\left(1-\frac{\sin^2\theta}{2(m^2+\kappa^2)}\right) \end{aligned}\right\} \quad . \quad . \quad (72),$$

showing that as a first approximation, and especially when the angle of incidence is small, m and κ may be taken to be equal to ν_0 and κ_0.

156. The Optical Constants of Metals. We owe to Drude† the best determination of the optical constants of metals After a careful investigation of the effects of the condition of the surface and the reflexion of surface films, results were obtained which are reproduced in Table XI. The measurements refer to sodium light

I have added the third and fourth columns giving the values of $\kappa^2 - \nu^2$ and $\nu\kappa\cos\rho$, the two invariants of metallic refraction The column headed θ_p gives the angle of principal incidence, the last column, the calculated reflected intensity for normal incidence

The table shows the remarkable fact that $\nu^2 - \kappa^2$ is negative for all metals. In the older theories of refraction in which the sympathetic vibrations within the molecule were neglected, this appeared to be an anomaly, for turning to equation (38), $\nu^2 - \kappa^2$ could not be negative independently of A and B unless κ were negative, which has no meaning.

* *Optik* p. 187 and *Wied. Ann* Vol. vii p 119 (1879)

† *Wied. Ann.* xxxix p 481.

TABLE XI

	κ_0	ν_0	$\kappa^2-\nu^2$	$\nu\kappa\cos\rho$	θ_p	$\left(\frac{\nu_0-1}{\nu_0+1}\right)^2$
Bismuth	3 66	1 90	9 79	6 955	77° 3′	652
Lead	3 48	2 01	8 07	6 995	76° 42′	621
Mercury	4 96	1 73	21 61	8 580	79° 34′	784
Platinum	4 26	2 06	13 90	8 776	78° 30′	701
Gold	2 82	0 37	7 81	1 043	72° 18′	851
Antimony	4 94	3 04	15 16	15 02	80° 26′	701
Tin	5 25	1 48	25·37	7 771	79° 57′	825
Cadmium	5 01	1 13	23 82	5 661	79° 22′	847
Silver	3 67	0 18	13 44	6607	75° 42′	953
Zinc	5 48	2 12	25 54	11 62	80° 35′	786
Copper	2 62	0 64	6 45	1 677	71° 35′	732
Nickel	3 32	1 79	7 82	5 943	76° 1′	·620
Cobalt	4 03	2 12	11 75	8 543	78° 5′	675
Steel	3 40	2 41	5 75	8 194	77° 3′	585
Aluminium	5 23	1 44	25 28	7 531	79° 55′	·827
Magnesium	4 42	0 37	19 40	1 635	77° 57′	929

TABLE XII.

Wave-lengths in tenth-metres	4500	5000	5500	6000	6500	7000
A *Pure Metals*	%	%	%	%	%	%
Platinum	55 8	58 4	61 1	64 2	66 3	70 1
Gold	36 8	47 3	74 7	85 6	88 2	92 3
Silver	90 6	91 8	92 5	93 0	93·6	94·6
Copper	48 8	53 3	59 5	83·5	89·0	90 7
Nickel	58·5	60 8	62 6	64 9	65 9	69·8
Steel (hard)	58 6	59 6	59 4	60 0	60 1	60 7
Steel (soft)	56 3	55·2	55·1	56 0	56 9	59·3
B *Speculum Metals*						
Alloy of Brashear 68·2% Cu+31 8 Sn	61 9	63 3	61 0	64 4	65·4	68 5
Alloy of Brandes and Schimemann. 41 % Cu + 26 Ni + 24 Sn + 8 Fe + 1 Sb	49 1	49·3	48 3	47 5	49 7	54·9
Alloy of L Mach $66\frac{2}{3}$ % Al + $33\frac{1}{3}$ Mg	83 4	83 3	82 7	83 0	82 1	83·3
C. *Glass Plates* covered at the back with						
Silver	79 3	81·5	82 5	82·5	83·5	84·5
	85 7	86·6	88 2	88·1	89·1	89·6
Mercury	72·8	70 9	71 2	69 9	71·5	72·8

The somewhat important question relating to the amount of light reflected at normal incidence has been investigated directly by E Hagen and H Rubens* Some of their results are embodied in Table XII

A comparison with Drude's numbers shows generally a good agreement. The exceptions are Platinum and especially Copper. The alloy of Brandes and Schunemann is of practical importance owing to its permanence and resistance to deterioration when exposed to moist or impure air

Drude has also determined ν_0 for red light and found that all metals except lead, gold, and copper, refract the red more than the yellow light The coefficient of extinction was determined directly by Rathenau †.

There is another difficulty which consists in the fact, already pointed out by Maxwell in his Treatise, that the simple theory according to which $\nu_0\kappa_0$ should be (neglecting B) equal to $CV\lambda$ gives too large values for the coefficient of extinction. The introduction of A and B does not entirely remove the difficulty, but as was first shown by Drude, a satisfactory explanation may be arrived at, if we adopt the theory of electrons throughout

The whole electric current according to this theory is carried by ions which possess effective inertia inasmuch as they possess energy proportional to the square of the velocity, the energy being that of the magnetic field they create. In the ordinary theory of conductors, this energy is only partially taken account of, the *average* magnetic field established by the moving electron being used in the calculations. If instead of an evenly distributed electric fluid, we imagine electricity to be concentrated within the electron, the magnetic field in the immediate neighbourhood of that electron will be very much larger than the average energy in each element of volume. Hence the inertia, or the coefficient of self-induction, by whichever name the factor in question may be called, is underrated in the ordinary treatment The error committed depends on the nearness of the moving electrons, and on their linear dimensions, but if their distance apart is great compared with their diameter, the additional energy per unit volume is $\frac{1}{2}\sigma i^2$ where i is the current density and σ stands for ρ/Ne^2 ‡, N being the number of moving electrons per unit volume, and ρ the apparent mass. If each molecule supplies one electron, which carries the conduction current, σ is of the order of magnitude 5×10^{-11} and has the dimensions of a surface

* Drude's *Annalen*, vol I p 352

† Quoted in Winkelmann, *Encyclopædia der Wissenschaften*, vol II p. 838

‡ A Schuster, *Phil Mag* vol I p 227 (1901)

The effect of the inertia is the same as that of an electric force $\sigma \frac{di}{dt}$ opposing the current If C be the conductivity this electric force produces a current density $C\sigma \frac{di}{dt}$ where i only relates to the conduction current

The equation of electric current (33) now becomes, if we denote by w' the z component of the conduction current,

$$w + C\sigma \frac{dw'}{dt} = \frac{K}{4\pi} \frac{dR}{dt} + CR + Ne\zeta,$$

also

$$w' = w - \frac{K}{4\pi} \frac{dR}{dt} - Ne\zeta$$

Hence if D stands as a symbol for $\frac{d}{dt}$

$$w + C\sigma \frac{dw}{dt} - CR + (1 + \sigma CD) \left(\frac{K}{4\pi} \frac{dR}{dt} + Ne\dot{\zeta}\right),$$

or with the help of (34), omitting the last term on the right-hand side of that equation,

$$w = CR + (1 + \sigma CD) \left(\frac{K}{4\pi} \frac{dR}{dt} + Ne\zeta\right) - \frac{\sigma C}{4\pi} \nabla^2 R$$

Differentiating with respect to the time and applying (34) again, we find

$$D \left\{4\pi CR + (1 + \sigma CD) \left(K \frac{dR}{dt} + 4\pi Ne\zeta\right)\right\} = (1 + \sigma CD) \nabla^2 R \quad \ldots (73).$$

If the motion is periodic and contains $e^{-\iota\omega t}$ as factor, we may substitute $-i\omega$ for D, and for ζ we may use its equivalent (37) in terms of R. The equation then becomes

$$\nabla^2 R = -(E\omega^2 + iF\omega) R \quad \ldots (74),$$

where

$$E = K + NeA - \frac{4\pi\sigma C^2}{1 + \sigma^2 C^2 \omega^2},$$

$$F = NeB\omega + \frac{4\pi C}{(1 + \sigma^2 C^2 \omega^2)}.$$

Hence from (13)

$$\left.\begin{aligned} \nu_0^2 - \kappa_0^2 &= V^2 \left(K + NAe - \frac{4\pi\sigma C^2}{1 + \sigma^2 C^2 \omega^2}\right) \\ 2\nu_0\kappa_0 &= V^2 \left(NBe + \frac{2C\tau}{1 + \sigma^2 C^2 \omega^2}\right) \end{aligned}\right\} \quad \ldots (75),$$

where in the last term, τ is written for $2\pi/\omega$

The effect of σ is therefore to diminish the product $\nu_0\kappa_0$ and hence to diminish that part of the absorption which depends on conductivity.

Equations (75) are, allowing for a change of notation, identical with those obtained by Drude. If we disregard A and B, we obtain

$$\left.\begin{aligned} \nu_0^2 - \kappa_0^2 &= 1 - \frac{4\pi\sigma C^2 V^2}{1 + \sigma^2 C^2 \omega^2} \\ \nu_0 \kappa_0 &= \frac{C V \lambda}{1 + \sigma^2 C^2 \omega^2} \end{aligned}\right\} \quad \ldots\ldots \quad (76).$$

If the numerical values of C and ω are introduced and the quantity σ is estimated, it is found that $\sigma C\omega$ is large, and equations (75) become with sufficient accuracy

$$\nu_0^2 - \kappa_0^2 = 1 - \frac{4\pi V^2}{\sigma\omega^2} + NAV^2 e,$$

$$2\nu_0\kappa_0 = V^2\left(\frac{2\tau}{\sigma^2 C\omega^2} + NBe\right).$$

As B is always positive, it follows that

$$\nu_0\kappa_0 > \frac{V\lambda}{C\omega^2\sigma^2}.$$

This relation allows us to calculate the number of electrons which take part in conduction currents, and it is found that this number in the different metals is of the same order of magnitude as the number of molecules*.

157. Reflecting powers of metals for waves of low frequency. Maxwell's theory has received an important confirmation in the work recently published by Hagen and Rubens†, on the relation between the optical and electrical qualities of metals. These investigations relate to waves of low frequency.

Neglecting σ, we may write equations (76)

$$\nu_0^2 - \kappa_0^2 = 1,$$

$$\nu_0\kappa_0 = C V\lambda.$$

As λ is supposed to be large, both ν and κ must be large and nearly equal. The second equation gives, neglecting the difference between the two quantities,

$$\kappa_0 = \sqrt{C V\lambda}$$

To test this formula for long waves, Hagen and Rubens measure the reflecting powers at normal incidence For the intensity of the reflected light, we have obtained the expression (64), which by substitution becomes $(\kappa_0 - 1)/(\kappa_0 + 1)$ for large values of κ_0.

Writing R' for the reflecting powers of a metal in per cent., $100 - R'$ gives the intensity of the light which enters the metal, the

* Schuster, *Phil Mag*, Vol VII p. 151 (1904)

† *Ann. d. Physik*, Vol. XI. p. 873 (1903) and *Phil. Mag* Vol. VII. p. 157 (1904).

intensity of incident light being 100, and the formula to be verified becomes

$$100 - R' = \frac{200}{\sqrt{C\tau}},$$

where τ is the period

TABLE XIII

1	2	3	4	5	6
Metals	$(100 - R')$ for $\lambda = 12\,\mu$		Conductivity at 170°	$(100 - R')$ for $\lambda = 25\ 5\,\mu$ & 170°	
	Observed	Computed		Observed	Computed
Silver	1 15	1 3	39 2	1 13	1 15
Copper	1 6	1 4	32 5	1 17	1 27
Gold	2 1	1 6	27 2	1 56	1 39
Aluminium			20 4	1 97	1 60
Zinc			10 2	2 27	2 27
Cadmium			8 40	2 55	2 53
Platinum	3 5	3 5	5 98	2 82	2 96
Nickel	4 1	3 6	5 26	3 20	3 16
Tin			5 01	3 27	3 23
Steel	4 9	4 7	3 30	3 66	3 99
Mercury			0 916†	7 66	7 55
Bismuth	17 8	11 5	0 513	25 6	10 09
Rotguss*			7 05	2 70	2 73
Manganin			2 37	4 63	4 69
Constantin	6 0	7 4	2 04	5 20	5 05
Patent Nickel P	5 7	5 4	3 69	4 05	3 77
Patent Nickel M	7 0	6 2	2 86	4 45	4 28
Rosse's Alloy	7 1	7 3			
Brandes and Schunemann's Alloy	9 1	8 6			

In Table XIII, columns 2, 3, the observed and computed values of $100 - R'$ are tabulated for $\lambda = 12\mu$ ($\mu = 10^{-4}$ cms.) For larger values of λ, R' approaches 100 % asymptotically with increasing wave-lengths, and the difficulty of experimentally determining $100 - R'$ increases accordingly. Hagen and Rubens therefore measured the emissive power instead of the reflecting power of the metals

From a comparison of the radiation sent out by a metal with that sent out by a black body, the reflecting power may be directly deduced.

* "Rotguss" contains 85 7 Cu + 7 2 Zn + 6 4 Sn

† At 100°

In Table XIII, column 4 gives the conductivity at 170° C. calculated from the known conductivities at 18° and the temperature coefficient. Columns 5 and 6 give the observed and computed emissive powers at 170° It will be noticed that the agreement is excellent in all cases except Aluminium which shows a considerable deviation, and Bismuth which forms a complete exception to the law

Professors Hagen and Rubens have also directly verified the fact that the quantity $100 - R'$ indicates with increasing temperature a change corresponding to the change of electrical resistance, and they further point out the remarkable fact that it would be possible to undertake absolute determinations of electrical resistance solely by the aid of measurements on radiation. The agreement of Maxwell's theory in its simple original form with the result of the experiments just described, proves that for wave-lengths as great as 12μ, the free periods of vibration of the molecules do not affect the optical constants of metals

158. Connexion between refractive index and density. The investigations of Sellmeyer and those based on similar principles all lead to the conclusion that for any one kind of molecule, $\mu^2 - 1$ is proportional to the density We possess a good many experimental investigations on the changes observed when the density (D) is altered by pressure or by temperature, and these investigations have not been favourable to the constancy of $(\mu^2 - 1)/D$

The failure of the formula in the case of a change from the liquid to the gaseous state is not perhaps surprising because the molecule in the liquid state may be expected to be more complex and behave optically very differently from the molecule of the same substance in the state of a gas. Even changes of temperature and pressure may to some extent affect the absorbing power, and consequently the velocity of propagation of light-waves, so that our theory cannot pretend completely to include such changes But though the failure of the theory is not surprising it must be pointed out that other formulae give better results. The simpler relationship which asserts the constancy of $(\mu - 1)/D$ discovered empirically by Gladstone and Dale, has often been successfully applied, and two authors of similar name, H A. Lorentz of Leyden*, and L Lorenz† of Kopenhagen, have almost simultaneously published investigations leading to the result that $(\mu^2 - 1)/(\mu^2 + 2)D$ is constant. The latter formula is capable of predicting with fair accuracy the refractive index of a gas, that of the liquid being known

Lord Rayleigh's investigations on the effect of spherical obstacles‡

* *Wied. Ann* IX p 641 (1880)

† *Wied. Ann* XI p 70 (1880)

‡ *Collected Works*, Vol IV p 19.

on the propagation of sound or light are very instructive, because they show clearly the various circumstances that may affect the problem. It seems clear that any investigation based on the effect of the influence of the presence of the molecules on the potential and kinetic energies must make $\mu^2 - 1$ and not $\mu - 1$ a function of the density. The investigations treated of in this Chapter belong to this group because our explanation of dispersion has introduced terms affecting the kinetic energy of the medium The investigations which lead to such expressions as $\left(\frac{\mu^2 - 1}{\mu^2 + 2}\right)/D$ neglect the vibrations of the molecules but consider their linear dimensions The molecules are supposed to be regularly spaced, occupying a volume not negligible compared with the total space, and to be made up of some material having a dielectric constant different from that of the surrounding space I think the experimental facts of selective refraction are sufficient to show that in the region of the spectrum which contains the free vibrations of the molecules, these constitute the paramount factor, but it is quite possible that when we wish to include the changes of molecular distance, which are brought about by changes of pressure or temperature, the linear dimensions must be taken into account.

There is another method of treating the subject which consists in considering the effect of a number of *irregularly* spaced molecules. This has been adopted by Lord Rayleigh in his investigations on the scattering of light by small particles The effect on the primary wave of each particle, whether due to sympathetic vibration or to a change in the optical properties, is as a first approximation a change of phase which when the molecular distance is small compared with a wave-length is equivalent to a change in wave velocity. By this reasoning we are led to the conclusion that $\mu - 1$ is proportional to the density When the total effect is small, $\mu - 1$ is proportional to $\mu^2 - 1$, so that as a first approximation, there is no difference between the two results, but there is a fundamental difference between the assumptions made in the two cases. If the distribution is regular, we need only consider the average kinetic and potential energies and may from them proceed to calculate the velocity of wave propagation, assuming that the progressive wave contains all the energy But with an irregular distribution there is a scattering of light in all directions, and consequent loss of energy The occurrence of $\mu^2 - 1$ in one case, and of $\mu - 1$ in the other, seems to be due to this distinction in the adjustment of energy, but when the molecules are as close as they are in solids and liquids, there must be considerable approach to regularity, and the scattering must be comparatively small. Hence theoretical considerations are in favour of μ^2

On the other hand Gladstone and Dale's empirical formula depending on $\mu - 1$ has had very substantial successes in the investigation of the effects of molecular composition on refractive power. The effect of the molecules being proportional to their number per unit volume, which is proportional to d/P if d is the density and P the molecular weight, the quantity $(\mu - 1) P/d$ was introduced by Landolt and called the molecular equivalent of refraction It was found that if in any compound there are n_1 atoms of one kind, n_2 of another, n_3 of a third, the molecular equivalent of refraction could approximately be calculated from the equation

$$\frac{(\mu - 1) P}{d} = n_1 v_1 + n_2 v_2 + n_3 v_3,$$

where v_1, v_2, v_3 are quantities which belong to the element.

159. Historical. Augustin Louis Cauchy, whose work has already been referred to at the end of Chapter x, published some important researches in wave propagation and first obtained formulae giving the constants of elliptic polarization of light on reflexion from metallic surfaces These he published however without proof.

Jules Célestin Jamin (born May 30, 1818, in the Department of the Ardennes, died February 12, 1886, at Paris) was the pioneer in the experimental investigation of metallic reflexion, and showed that Cauchy's equations represented the facts with sufficient accuracy. Eisenlohr supplied the analytical proof of Cauchy's formulae and showed that the proper interpretation of Jamin's measurements leads to the conclusion that for silver, the refractive index is smaller than one. This result, which did not seem at that time to be reconcilable with the stability of the medium inside the metal, received support from Quincke's experiments which proved an *acceleration* of phase when light passed through thin metallic films The matter was finally settled by A Kundt* (born Nov. 18, 1839 at Schwerin, died May 21, 1894 near Lubeck, Professor of Physics in the University of Berlin) who succeeded in making thin prisms of metals and thus could demonstrate directly that in metals light was propagated more quickly than in vacuo The apparent anomaly of this result received its explanation when the refraction of absorbing media generally was more carefully studied

In 1862 Le Roux having filled a hollow prism with the vapour of iodine, noticed that while it absorbed the central parts of the spectrum, it transmitted the red and violet ends, refracting however the red end more than the violet This phenomenon he called anomalous dispersion. Eight years later, Christiansen noticed the same phenomenon in the

* *Wied. Ann* xxxiv p 469 (1888)

case of a solution of fuchsin The matter then attracted considerable attention, and A Kundt especially improved the experimental methods. Including a great many colouring matters in his investigations he was able to formulate the general laws which regulate the influence of absorption on refraction In the meantime, Sellmeyer had published his theoretical investigation, which is now generally recognized to be correct in principle It only remains to allude to the work of Ketteler, who more than any one else has shown, both by experiment and by mathematical calculation, that all refraction is of one kind, and that even in the case of apparently transparent media like water, it is necessary to take account of the effects of the free vibrations of the molecules both in the infra-red and ultra-violet

The recent development of the subject has already been sufficiently treated

CHAPTER XII.

ROTATORY EFFECTS

160. Photo-gyration In all cases hitherto considered the transmission of a luminous disturbance has been such that a plane polarized wave was propagated with its plane of polarization remaining parallel to itself But there are media in which the wave, though remaining plane polarized, shows a continuous rotation of the plane of polarization as it proceeds If plane polarized light be made to traverse, for instance, a tube filled with a sugar solution, and the emergent light be examined, it is observed that the plane of polarization has been turned through an angle which is proportional to the length of the tube and also depends on the concentration of the solution

The direction of rotation is different in different substances. It is said to be right-handed when it is in the direction of the rotation of the hands of a watch, looked at from the side towards which the light travels

Substances which possess this property are often called "optically active," an expression which is not very descriptive and possibly misleading, as the word "activity" has been applied to several different properties We shall find that the distinctive feature of the rotational property is the different velocity of propagation of circularly polarized light according as it is right-handed or left-handed We may therefore appropriately call substances "dextrogyric" or "laevogyric" according as they turn the plane of polarization to the right or to the left (γῦρος, a circle) A substance is simply called "photogyric" if it acts in its isotropic state, but "crystallogyric" if, like quartz, the property is connected with its crystalline nature Finally all substances turn the plane of polarization when they are traversed by light in the direction of a magnetic field They become therefore "magneto-gyric" If a special word be required to express the general property not applied to any particular case, we shall use the expression "allogyric" (ἄλλος, different), while substances which are inactive are "isogyric"

The allogyric property implies some asymmetrical structure, and in the case of solutions, the want of symmetry must be in the structure of the molecule itself. Van 't Hoff and le Bel have indeed drawn important conclusions as to the arrangement of the atoms in the molecule of allogyric substances

Quartz is the most conspicuous example of a crystallogyric body. If a plate a few millimetres thick be cut out of a crystal of quartz perpendicularly to the axis, and this plate be examined between crossed Nicols, the luminosity of the field is seen to be restored If the original light was white, the transmitted light is coloured The explanation of the effect presents no difficulty on the assumption of a rotation of a plane of polarization which is different for different wave-lengths There is no rotation of the plane of polarization if the wave-front is parallel to the axis Some specimens of quartz show a right-handed rotatory effect while others are left-handed It is found that generally the direction of the rotation may be detected by a close examination of the crystal, there being certain small asymmetrical planes at the corners between the hexagonal prism and pyramid, the position of which is different for the two types of crystals In all substances hitherto discovered, which are allogyric, the angle of rotation per unit length of substance traversed increases with the refrangibility and is approximately proportional to the inverse square of the wave-length

There is a marked distinction between the magnetogyric and other allogyric effects In the case of substances which possess the rotatory property in their natural state, the rotation for rays travelling opposite ways is in the same direction when looked at from the same position *relative to the direction in which the light travels* Thus if A and B are two ends of a tube containing a solution of sugar and light is sent through the tube from A to B, an observer looking at B towards the light will observe a right-handed rotation If now the light be sent from B to A and the observer looks at A, the rotation observed by him is still right-handed. If there were a mirror at B, and the ray after traversing the tube from A to B were reflected back towards A, the plane of polarization at emergence would be parallel to the direction it had on first entering the tube at A This we should indeed expect by the principle of reversibility (Art 25). In the case of magnetogyration on the contrary the direction of rotation is different as seen by the observer according as the light travels with or against a line of force, but it is the same when looked at from the same position *relative to the direction of the magnetic field*. Consequently if light travels from A to B, and is reflected back at B, the angle through which the plane of polarization is rotated is increased and finally doubled during the

passage backwards The principle of reversibility holds in this case also, provided we reverse the direction of the magnetic field as well as the direction of the ray

161 Analytical representation of the rotation of the plane of polarization. Consider plane waves travelling in the direction of x, with a uniformly rotating direction of vibration As each wave-front reaches a given position, the direction of vibration is a definite one, and the angle which that direction forms with one fixed in space is therefore a function of x only If it be a linear function of x, the plane of polarization rotates through an angle which is proportional to the distance traversed Let η and ζ be the projections of the displacement, and put

$$\left.\begin{aligned} \eta &= 2\cos rx \cos(lx-\omega t) \\ \zeta &= 2\sin rx \cos(lx-\omega t) \end{aligned}\right\} \quad . \qquad (1)$$

The equations satisfy the condition laid down for the direction of vibration, for if δ be the angle between it and the axis of z

$$\begin{aligned} \tan\delta &= \frac{\eta}{\zeta} \\ &= \tan rx, \end{aligned}$$

from which it follows that δ is a linear function of x, and that r measures the angle of rotation per unit length of path We call the quantity r the "gyric coefficient" Equations (1) also satisfy the conditions of ordinary wave propagation, as the displacements may be expressed as a sum of terms, each of which has the form $f(x-vt)$ To show this we need only transform the products of the circular functions in a well-known manner

Writing

$$\left.\begin{aligned} \eta_1 &= \cos(l_1x-\omega t), \quad & \eta_2 &= \cos(l_2x-\omega t) \\ \zeta_1 &= \sin(l_1x-\omega t), \quad & \zeta_2 &= -\sin(l_2x-\omega t) \end{aligned}\right\} \qquad (2),$$

we find that (1) becomes identical with

$$\eta = \eta_1 + \eta_2, \qquad \zeta = \zeta_1 + \zeta_2,$$

provided that

$$r = \tfrac{1}{2}(l_1 - l_2), \quad l = \tfrac{1}{2}(l_1 + l_2),$$

or if r and l be given

$$l_1 = l + r; \quad l_2 = l - r.$$

The disturbance is now expressed in terms of four parts, each of which is of the homogeneous type, but while the periodic time for each of these four waves is the same, the wave-lengths are in groups of two · $2\pi/l_1$ and $2\pi/l_2$ respectively. The displacements η_1 and ζ_1 form together a right-handed circularly polarized ray, propagated with velocity $v_r = \omega/(l+r)$, while the displacements η_2 and ζ_2 combine to

form a left-handed circularly polarized ray propagated with velocity $v_l = \omega/(l-r)$ The gyric coefficient may be deduced from v_r and v_l by means of

$$r = \frac{\omega}{2}\left(\frac{1}{v_r} - \frac{1}{v_l}\right). \qquad (3)$$

The important conclusion that a wave travelling with a uniform rotation of its planes of polarization is equivalent analytically to the superposition of two circularly polarized rays of opposite directions and propagated with different velocities is due to Fresnel A simple geometrical illustration may be given If two points P and Q are imagined to revolve in opposite directions with uniform and identical velocities round the circumference of a circle (Fig. 171), they will cross at two opposite ends A and B of a diameter, and their combined motion is equivalent to a simple periodic motion along AB as diameter. The two points may be considered to represent the displacements of two waves polarized circularly in opposite directions, having for their resultant a plane polarized wave. If the two circularly polarized waves are transmitted with different velocities, there is, as the waves proceed, a gradual retardation of one circular motion relative to the other, so that the crossing points gradually shift to one side or the other The combined motion always remains a simple periodic motion along a diameter, but that diameter rotates uniformly as we proceed along the wave normal If A_1, B_1 are the crossing points in a wave-front which is at unit distance from that originally considered, AOA_1 represents the angle through which the plane of polarization is turned in unit length of path

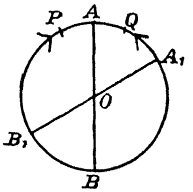

Fig 171

162. Isotropic Substances There is no satisfactory representation of the mechanism by means of which an asymmetrical molecular structure turns the plane of polarization, but we may easily extend our former equation so as to include rotatory effects. Our equation (41) Chapter XI., for the displacement of an electron,

$$\ddot{\zeta} + n^2\zeta = \frac{eR}{\rho} \qquad \ldots \ldots (4),$$

assumed that the electron suffers no constraint in its motion, but if other forces act which depend on the displacements of other electrons, the resultant force may involve not only the three components P, Q, R of electric force, but also their nine differential coefficients with respect to the three independent space variables Considering small motions only, we need only take linear terms into

account The complicated general equation which would result from the substitution on the right-hand side of (4) of twelve linear terms is much simplified by the restriction that our investigation shall only apply to isotropic substances

In such substances a luminous wave is affected equally in whatever direction it passes, and the resultant differential equation must therefore be independent of the direction of the coordinate axes If for instance we turn the system of axes through 180° round the axis of z, the simultaneous reversal of the signs of P, Q, x and y must leave the equations unaltered This consideration shows that there are no terms involving Q, R, $\frac{dP}{dz}$, $\frac{dQ}{dz}$, $\frac{dR}{dx}$ and $\frac{dR}{dy}$ because all these terms if existing would reverse their sign by the supposed change in the coordinate axes Similarly if we rotate the axes through 180° round the axis of x, the left-hand side of (4) changes sign, hence the general term to be substituted on the right must also reverse its sign This excludes the terms depending on $\frac{dP}{dx}$, $\frac{dQ}{dy}$, $\frac{dR}{dz}$. The only remaining differential coefficients are $\frac{dP}{dy}$ and $\frac{dQ}{dx}$, and these must occur in the combination $\frac{dP}{dy} - \frac{dQ}{dx}$, as may be seen by turning the system through 90° round the axis of z and introducing the condition that the equation remains unaltered We may therefore write the resulting differential equation

$$\ddot{\zeta} + n^2\zeta = \frac{e}{\rho}\left[R + s\left(\frac{dP}{dy} - \frac{dQ}{dx}\right)\right] \qquad (5)$$

Similarly
$$\ddot{\eta} + n^2\eta = \frac{e}{\rho}\left[Q + s\left(\frac{dR}{dx} - \frac{dP}{dz}\right)\right] \qquad (6)$$

Confining ourselves to insulators, equation (35) of Chapter XI is

$$K\frac{d^2R}{dt^2} = \nabla^2 R - 4\pi Ne\zeta \qquad (7)$$

If the displacements are proportional to $e^{-\iota\omega t}$, so that in (5) we may write $-\ddot{\zeta}/\omega^2$ for ζ, we obtain by substitution in (7)

Similarly
$$\left.\begin{aligned} K\frac{d^2R}{dt^2} &= \nabla^2 R - \omega^2 m\left[R + s\left(\frac{dP}{dy} - \frac{dQ}{dx}\right)\right] \\ K\frac{d^2Q}{dt^2} &= \nabla^2 Q - \omega^2 m\left[Q + s\left(\frac{dR}{dx} - \frac{dP}{dz}\right)\right] \end{aligned}\right\} \qquad (8),$$

where
$$m = \frac{4\pi Ne^2}{\rho(\omega^2 - n^2)}$$

If we consider plane waves parallel to yz so that the electric forces are independent of y and z, equations (8) may be written

$$\left.\begin{aligned}\frac{d^2R}{dx^2} &= -\left(M_1R - M_2\frac{dQ}{dx}\right)\\ \frac{d^2Q}{dx^2} &= -\left(M_1Q + M_2\frac{dR}{dx}\right)\end{aligned}\right\} \quad . \quad . \quad . \qquad (9),$$

where $$M_1 = \omega^2(K+m), \quad M_2 = \omega^2 ms$$

From equations (9) we derive, if $i = \sqrt{-1}$

$$\frac{d^2}{dx^2}(Q+iR) = -\left[M_1 - iM_2\frac{d}{dx}(Q+iR)\right] \qquad (10)$$

This has for a particular solution

$$Q + iR = e^{i(l_1x - \omega t)} \quad . \quad .. \qquad (11),$$

provided that

$$l_1^2 = M_1 + M_2l_1 \quad . \quad . \quad . \qquad (12)$$

Reversing the sign of i in (10) and assuming a solution

$$Q - iR = e^{i(l_2x - \omega t)} \quad . \qquad (13),$$

we find the equation of condition

$$l_2^2 = M_1 - M_2l_2 \quad . \quad . \qquad (14)$$

The positive roots of (12) and (14) which alone need be considered are

$$\left.\begin{aligned}2l_1 &= M_2 + \sqrt{M_2^2 + 4M_1}\\ 2l_2 &= -M_2 + \sqrt{M_2^2 + 4M_1}\end{aligned}\right\} \qquad (15).$$

Separating and retaining only the real parts in the solutions (11) and (13), it is seen by comparison with (2) that (11) represents a right-handed circular polarization, while (12) represents a left-handed circular polarization The two waves are propagated with velocities ω/l_1 and ω/l_2 respectively The superposition of both solutions represents a plane polarized wave, the plane of polarization rotating per unit length of optical path through an angle r which is obtained from (3)

$$\begin{aligned} r &= \tfrac{1}{2}(l_1 - l_2) = \tfrac{1}{2}M_2\\ &= -\frac{2\pi Ne^2s}{\rho}\frac{\omega^2}{\omega^2 - n^2} \quad . \quad ... \quad . \quad .. \qquad (16)\end{aligned}$$

The above investigation shows that the only terms depending on the first differential coefficients of the electric forces which can be added to the general equations of light and are consistent with isotropy indicate a turn of the plane of polarization. This does not of course furnish an explanation of the rotatory effect, which would require a knowledge of the physical cause for the existence of the terms We may however take one step forward towards an explanation by

considering that the terms in equation (5) which have been added represent a torsional electric force having the axis of ζ as axis The equations mean therefore that a displacement of the electron in the z direction may be produced not only by a force acting in that direction, but also by a couple acting round it A rifle bullet lying in its rifle barrel would be displaced in a similar manner along the barrel both by a pulling and twisting force But if we take the dimensions of a single electron to be very small, we exclude the possibility of a constraint which would enable a couple to cause a motion in one direction. We must in that case draw the conclusion that the vibrations of the electron which give rise to the rotatory effect are motions of systems of electrons united together by certain forces which are such that a couple of electric forces produces a displacement of the positive electrons in one direction or of the negative electrons in the opposite direction along the axis of the couple In view of the fact that a single electron cannot be acted on by a torsional force, it would have been more appropriate to base our investigation on equation (44) Chapter XI. The generalized force Ψ_1 would in the present problem depend not only on R but on $\left(\frac{dP}{dy} - \frac{dQ}{dx}\right)$, and if the investigation in the second part of Art 149 is modified by the addition of appropriate terms, the result arrived at would, for a single variable, remain the same as that represented by (16)

163. Allogyric Double Refraction. Equations (2) show that the analytical representation of plane polarized waves travelling through an optically active medium necessarily involves two different wave velocities In any question concerning the refraction and reflexion of light, we may take all four displacements represented by (2) separately and apply the formula obtained for homogeneous disturbances It is clear that the wave on emergence must be split into two separate waves which are circularly polarized in opposite directions. This double refraction, due to the rotational effect, is verified by experiment and has some practical importance. Quartz, as has already been mentioned, turns the plane of polarization of waves travelling parallel to the optic axis, and in consequence, a ray travelling along the optic axis is doubly refracted at emergence Quartz is very useful in optical investigations on account of its transparence to ultra-violet rays, and it is a serious drawback that it is impossible to avoid double refraction in a prism made of that substance. The difficulty is overcome by combining two prisms made of two specimens, one of which has a right-handed and the other a left-handed rotatory power These two prisms ABC and A_1BC (Fig 172) are right-angled at C and have their optic axes parallel to AA_1 They

are joined together along BC, and if a ray traverses such a prism at minimum deviation its direction inside the prism is parallel to the axis. A ray polarized either in the principal plane or at right angles to it, is divided into two rays circularly polarized in opposite directions The same is true therefore of an unpolarized ray. Of these circularly polarized rays one gains over the other while traversing the first prism, and loses equally while passing through the second prism The combined optical distance is therefore the same for both components, and there is only a single refraction at emergence.

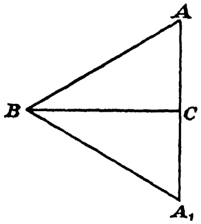

Fig 172

164. Crystalline Media The complete investigation of crystalline substances is complicated and serves no useful purpose, as rotatory effects have only been observed in uniaxal crystals We may therefore take quartz to be the typical example. Quartz shows no rotatory effects for rays travelling at right angles to the axis. A plane wave travelling in that substance splits up into two plane polarized waves if the wave travels at right angles to the axis, and into two circularly polarized waves if it travels parallel to the axis. In the case of waves travelling obliquely to the axis we may therefore surmise that the two waves are elliptically polarized, the ellipse becoming more and more eccentric as the wave becomes less inclined to the axis This conclusion is verified by experiment The elements of the ellipse have been made the subject of calculation by Sir George Airy*. A very clear account of the work of Airy, Jamin and Gouy on this subject is given by Mascart†

165. Rotatory Dispersion The rotation per unit length according to (16) is

$$r = \frac{\beta\omega^2}{n^2 - \omega^2},$$

on the supposition that we need only consider one period $2\pi/n$ of the free vibration In this expression β is a constant which can either be positive or negative according to the sign of s If the free period is very short compared with the range of visible periods, we may neglect ω in comparison with n, and the rotation is in that case proportional to ω^2, *i e.* inversely proportional to the square of the wave-length. This law holds approximately for most substances which have been examined. In general we have to consider several free periods, so that we must write

$$r = \Sigma \frac{\beta_m \omega^2}{n_m^2 - \omega^2} \quad \ldots\ldots\ldots\ldots (17),$$

* *Camb. Phil. Trans.*, Vol. IV. Part 1 (1831)

† *Optique*, Vol. II p. 513.

the summation having to be carried out for the different values of m. If the free periods lie in the ultra-violet so that all values of n_m are larger than ω we may expand the function in powers of ω and obtain after rearranging terms

$$r = r_1\omega^2 + r_2\omega^4 + r_3\omega^6 \quad \ldots \quad (18),$$

where r_1, r_2, r_3, are quantities depending on the values of β_m and n_m. The rotatory properties of quartz have been investigated over a very wide range It is found that the effects may be explained by assuming two ultra-violet free periods, one of which may be made to coincide with the ultra-violet period, which has been deduced from the general dispersion effects of quartz (Art. 150), the other being very short*. The infra-red periods necessary for the explanation of refraction do not seem to produce any rotatory effects

166 Isochromatic and Achromatic Lines The appearance of photogyric crystals in the polariscope is materially affected by their rotatory effect The calculation of the isochromatic and achromatic lines has been carried out by Sir George Airy A full account is given in Mascart's *Optics*† The simplest case is that of a plate cut at right angles to the axis examined with crossed polarizer and analyser Apart from the rotatory effect, the appearance should be that of Fig 1, Plate II Now owing to this rotatory effect the vibration which enters near the centre parallel to the principal plane of the polarizer leaves it inclined at an angle to that direction and is not therefore completely blocked out by the analyser The result is that there are no achromatic lines near the centre The general appearance is that of the figure, omitting the dark cross within the first dark ring

167 The Zeeman effect Before discussing the theory of photogyric effects, which a magnetic field impresses on a wave of light passing through it, we may give a short account of the modifications of the luminous radiations observed when the source of light is subjected to strong magnetic forces It was discovered by Zeeman in 1896 that a sodium flame placed in a magnetic field showed a widening of the two yellow lines, and at the suggestion of H A Lorentz, who at once foresaw the right explanation, further experiments were made to test the polarization of the emitted radiations which confirmed Lorentz's theory In the case of spectroscopic lines, which show the simplest type of magnetic effect, it is found that if the light is examined axially, *i e* parallel to the lines of force, each line splits into two, which are circularly polarized in opposite directions Looked at

* Drude, *Optik*, p 381

† *Optique*, Vol II p 314

equatorially, each line is divided into three components, the centre one being polarized in an equatorial plane, and the two others in a plane passing through the lines of force.

If we look upon the radiations as being due to the vibrations of an electron these observations admit of a simple explanation. Consider, first, light sent out in the axial direction. Each rectilinear vibration may be supposed to be made up of two opposite circular vibrations, the orbits lying in the equatorial plane. Let the light which reaches the observer travel through the flame in the direction of the lines of force, *i.e.* from the north to the south magnetic pole. A positive electron performing an anti-clockwise rotation, *i.e.* a positive rotation round a line of force, will under these circumstances be acted on by a force Hev, tending to increase the diameter of the circle in which it revolves (H = intensity of magnetic field, v = linear speed of electron, e = charge of electron). If in the absence of the magnetic field the acceleration is n^2d, where d is the displacement, and n_1^2d represents the acceleration in the circular path when the magnetic force acts, we have, ρ being the mass,

$$\rho n^2 d - Hev = \rho n_1^2 d.$$

Also $v = n\omega$ if ω is the time factor. Hence

$$\rho(n^2 - n_1^2) = \pm He\omega \qquad (19),$$

where the upper sign holds for the positive rotation.

In the case here considered we may write $\omega = n$, and considering $n - n_1$ always to be a small quantity, we may neglect its square. We thus find

$$n - n_1 = \frac{He}{2\rho}.$$

Finally, introducing the frequencies N and N_1 in place of n and n_1 we obtain

$$N - N_1 = \frac{He}{4\pi\rho},$$

or if we write

$$z = \frac{e}{4\pi\rho} \qquad (20),$$

$$N - N_1 = \pm zH \qquad (21).$$

The coefficient z may conveniently be called the Zeeman coefficient. We conclude that a rectilinear simply periodic motion is divided into two circular motions, the longer period showing anti-clockwise rotation if e is positive. Zeeman observed that the less refrangible component rotates clockwise, and the more refrangible one anti-clockwise, if the field is in the specified direction, and it follows that if our theory is correct it is the negative electron that gives rise to all vibrations for which this is the case.

Looked at equatorially, the two circular orbits appear in projection as rectilinear vibrations, and send out light vibrating at right angles to the lines of force So far the vibrations which take place along the lines of force have not been taken into account, but these are not in the simple theory here considered affected by the magnetic field. They constitute therefore plane polarized vibrations transmitted in the equatorial plane having an unmodified period Looked at equatorially we should expect therefore to see each line divided into three, the external components of the triplet having the same period as the circular vibrations observed in the axial direction. The agreement of the appearance reasoned out in this fashion with the observed facts constitutes a direct proof that the direction of vibration is at right angles to the plane of polarization, identifying variations in electric force with the direction of vibration Although there is much indirect evidence in favour of this view, such a convincing demonstration as that afforded by the Zeeman effect is very satisfactory

Our calculation has tacitly assumed that the vibrating electron is free from constraint and acts as an independent unit with three degrees of freedom If we dropped this assumption we should be led to more complicated magnetic effects, and indeed the majority of spectroscopic lines do not show the simple subdivision which the theory in its simplest form gives us

H A Lorentz* in a general theoretical discussion shows that if a spectroscopic line divides into n components, there must be n degrees of freedom in the system which in the absence of the magnetic field are coincident

According to the simplest form of the theory, the vibrations parallel to the lines of force preserve their period, but there are important cases in which these also change and two vibrations, one of larger and one of shorter period, take the place of the original one. In some cases the original period is maintained as well, in other cases it completely disappears Such a phenomenon shows that the vibration along the line of force is not free, but is accompanied by changes in directions at right angles to itself, and that the magnetogyric effect of the accompanying changes reacts on the original vibration.

In view of the importance of the subject, I give a short statement of some of the principal facts which have been established It is necessary to introduce it by a brief description of the structure of line spectra Many of the metallic spectra contain a number of lines which form a connected series, and we distinguish between the trunk series (Kayser and Runge's "Hauptserie"), the main-branch series (Kayser

* *Rapports présentés au congrès international de Physique de* 1900 Vol III. page 1

and Runge's "Zweite Nebenserie"), and the side-branch series (Kayser and Runge's "Erste Nebenserie") Figure 173 shows diagram-

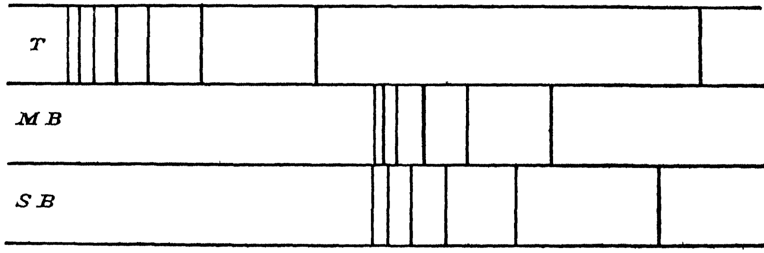

Fig 173.

matically the arrangement of lines in the three series, the red end of the spectrum being to the right, the trunk, main branch and side branch being marked T, $M B$, $S B$, respectively It is seen that the members of each series approach some definite limit of frequency on the more refrangible side; the point to which each converges I call the "root" of the series. The two branches have a common root at some point in the trunk According to an important law discovered by Rydberg, and later independently by the author of this book, the frequency of the common root of two branches is obtained by subtracting the frequency of the root of the trunk from that of its least refrangible member In the spectra of the alkali metals each line of the trunk is a doublet, and we may speak of a twin trunk springing out of the same root In the same spectra the lines belonging to the two branches are also doublets. The two components of the branch series are not twins springing out of the same root, but the difference in the vibration number is the same for each doublet, there being two roots giving the same difference. Rydberg's law being true for each component of a twin trunk, each set of components of the main branch is associated with one of the two divisions of the trunk, the root of *least* frequency being attached to that part of the trunk the members of which have the *highest* frequency The connexion of the side branches is not so clearly established, there being evidence of a further relation between them and hitherto unknown trunks

The lines which belong to the branches of the spectra of magnesium, calcium, and of the allied metals occur in triplets, and analogy leads us to think that the trunk must be a triplet also, each of the three compounds having one main and one side branch The trunk vibrations however have not been seen or identified in those metals.

Passing on to the behaviour of the different vibrations in the magnetic field, it was first announced by Preston, but more particularly proved

by Runge and Paschen, that each member of a trunk or of a branch behaves alike, not only as regards the general type of subdivision but also as regards the amount of separation in a given magnetic field, provided the separation is measured on the frequency scale The same is true for all series which correspond to each other Thus *M B* I II. and III, Fig 174, give the magnetic separation of the lines in each of the three main branches, in all cases where there are three main branches associated with each other (zinc, cadmium, mercury, calcium)

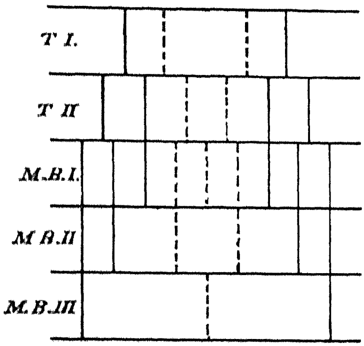

Fig. 174.

The figure represents the observations when the line of sight is at right angles to the magnetic field, each line being split up into the group shown in the figure The dotted lines represent the components vibrating parallel to the field, and the full lines those at right angles to the field. The portions of the figure marked *T* I and *T* II. represent the separation observed in the two sodium lines, and also apply in all cases where twin trunks exist (*e g* copper, silver), *T* I representing the type of the least refrangible component The same type is observed in a doublet found in each of the spectra of magnesium, calcium, strontium and barium (*e g* *H* and *K* of calcium), and we may therefore conclude that these doublets belong to a previously unknown trunk

The distances between the Zeeman components of each line are found by Runge and Paschen to be small multiples of a number, which is the same for each of the two members of the twin trunks. It is also the same for each of the main branches of the mercury and allied triplets. Thus in Fig. 174 the distances between the lines of the rows marked *T.* I and *T* II are all capable of being represented as small multiples of a difference in period which for the field used by Runge and Paschen (31000) was measured by them to be 0 459 (the unit here is the number of waves spread over one centimetre) Referred to the same scale and the same intensity of field, the common factor of the subdivision of the triplets marked *M B* I II. and III. is 0 702. Runge and Paschen point out that these numbers are very nearly in the ratio of 2 3 Comparing the types marked *T.* II and *M.B.* I. the figure shows that the smallest displacement of the vibrations at right angles to the field is the same in the case of the doublet and the triplet.

The types found in the twin trunk are also found in the main branches of the alkali metals with the difference that the type of the most refrangible member of the twin trunk is the same as the type of the least refrangible member of the branch doublet

The lines of the side branches of copper, silver, thallium and aluminium are doublets, the less refrangible member being accompanied by a satellite, which shows a complicated structure in the magnetic field Fig 175 shows the appearance ($S\,B.S$) of the satellite in the magnetic field, and also the types belonging to the least refrangible ($S\,B$ I) and most refrangible ($S\,B$ II) components of the side branches

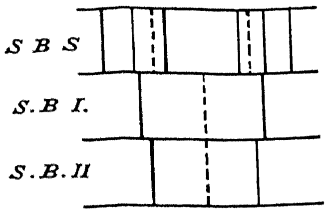

Fig. 175.

Much remains to be done in extending the investigation to other metals. Iron has been investigated pretty carefully Among the various types of separation Fig 176 shows three remarkable ones

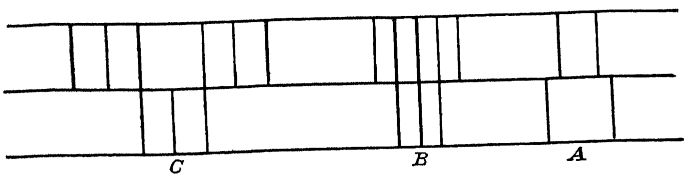

Fig 176

given by H. Becquerel and H Deslandres* In this figure the components vibrating perpendicularly are drawn above those which are parallel to the magnetic field The peculiarity of the type marked A is that the vibrations parallel to the field are more affected than those at right angles to it According to Berndt the green line of Helium is divided in accordance with this type A, but the experiment is difficult in the case of permanent gases and further measurements are much needed It will be noticed that in some cases the same component appears, whether the direction of vibration is at right angles to the field or parallel to it It would be interesting to notice whether in such cases the light is really elliptically polarized, as it should be if the coincidence were absolute

In Table XIV I have collected the calculated Zeeman coefficients for a few of the important types

* *C R.* CXXVII p 18 (1898).

TABLE XIV.

Type	Example. Wave-length in 10th metres	$z \times 10^{-5}$	$z' \times 10^{-5}$
DOUBLETS			
Trunk Series I	Na 5896	18 0=4×4 51	9 01 =2×4 52
Main Branch II	Al 3944		
Trunk Series II	Na 5890	22 0=5×4 50	
Main Branch I.	Al 3962	13 3=3×4 43	4 36
Side Branch I	Tl 3519	14 3=5×2 86	0
„ II	2768	11 2=4×2 80	0
Satellite	3530	19 8=7×2 83	10 4=4×2 60
		14 0=5×2 80	
		8 1=3×2 70	
TRIPLETS			
Main Branch I.	Zn 4811	27 0=4×6 75	6 79
	Hg 5461	20 5=3×6 83	0
		13 4=2×6 70	
„ II	Zn 4722	27 2=4×6 80	6 99
	Hg 4359	20 6=3×6 87	
„ III	Zn 4680	27 4=4×6 80	0
	Hg 4047		
Seven Mercury lines not belonging to any series		13 6	0
Two „ „ „		16 1	0
Two „ „ „		21 6	0

The members of the doublets and triplets are numbered in the order of diminishing wave-lengths, thus "Main Branch III" means the most refrangible member of a triplet belonging to the main branch series The table is entirely based on the measurements of Runge and Paschen* and the wave-lengths given in the second column as examples belong to lines from which, among others, these authors have derived their results. The last column gives under the heading z' the alteration in frequency (multiplied by 10^{-5}) of the vibrations which are parallel to the field Where several numbers are given in the third or fourth column, it means that the original vibration is split into more than two components

* *Berl Abh* (Anhang) 1902 *Sitzungsber d. Berl Ak* xix p 380 (1902) and xxxii p 720 (1902)

On the simple assumption under which equation (19) has been deduced, (20) allows us to calculate the important ratio e/ρ from the Zeeman coefficient. Taking z from Table XIV., the values of $4\pi z$ are found to range from $3{\cdot}4 \times 10^7$ to 10^7. Independent measurements of e/ρ founded on the properties of kathode rays, give values between 3 and $1{\cdot}86 \times 10^7$, the correct value lying probably nearer the higher than the lower of these numbers. The average apparent mass of the electron vibrating in the magnetic field calculated on the simple Lorentz theory is therefore not far different from that obtained by observation on kathode rays. There seem however, undoubted cases (probably the majority) where the constraint or mutual influence of the vibrating electrons *diminishes* their apparent mass. The simplest form of resolution in which each line, looked at transversely to the lines of force, is separated symmetrically into three components, of which the outer ones vibrate at right angles to the field, is found in the case of the side branch doublets of Copper, Silver, Aluminium, Thallium and Copper as well as in certain doublets of Calcium, Strontium, Magnesium and Barium. In all these cases the magnetic effects are identical when measured on the frequency scale. The values of e/ρ deduced from Runge and Paschen's measurements are $1{\cdot}8 \times 10^7$ and $1{\cdot}4 \times 10^7$ for the least and most refrangible members respectively. These numbers agree very well with those obtained by electrical measurement. On the other hand, the third main branch of the magnesium series, also showing the simplest resolution, gives a number about twice as great for e/ρ.

It is a significant fact that no Zeeman effect has yet been observed in the case of spectra of fluted bands such as those of carbon and nitrogen*. The magnetogyric properties of gases giving by absorption spectra of fluted bands render it very possible that such effects exist but have not been detected owing to their smallness. A slight increase in power may bring them to light.

168. Photo-gyration in the magnetic field. If an electron attracted to a fixed centre with a force varying as the distance moves in a magnetic field, its equations of motion are

$$\left.\begin{aligned}
\frac{d^2\xi}{dt^2} + n^2\xi &= \frac{e}{\rho}\left[P + \left(H_3\frac{d\eta}{dt} - H_2\frac{d\zeta}{dt}\right)\right]\\
\frac{d^2\eta}{dt^2} + n^2\eta &= \frac{e}{\rho}\left[Q + \left(H_1\frac{d\zeta}{dt} - H_3\frac{d\xi}{dt}\right)\right]\\
\frac{d^2\zeta}{dt^2} + n^2\zeta &= \frac{e}{\rho}\left[R + \left(H_2\frac{d\xi}{dt} - H_1\frac{d\eta}{dt}\right)\right]
\end{aligned}\right\} \quad \ldots\ldots\ldots (22),$$

where H_1, H_2, and H_3 are the components of magnetic induction due

* C. R. cxxvii. p. 18 (1898).

to the external field The right-hand side of the equations which express the components of electromagnetic force may easily be proved from the consideration that the force is at right angles both to the direction of the field and the direction of motion*

If we take the magnetic field to be of uniform strength H, the lines of force being parallel to the axis of x, the above equations may be written more simply

$$\left.\begin{aligned} \ddot{\xi} + n^2\xi &= \frac{e}{\rho} P \\ \ddot{\eta} + n^2\eta &= \frac{e}{\rho}(Q + H\dot{\zeta}) \\ \ddot{\zeta} + n^2\zeta &= \frac{e}{\rho}(R - H\dot{\eta}) \end{aligned}\right\} \quad \dots\dots \quad (23)$$

These equations together with

$$\left.\begin{aligned} K\ddot{P} &= \nabla^2 P - 4\pi Ne\ddot{\xi} \\ K\ddot{Q} &= \nabla^2 Q - 4\pi Ne\ddot{\eta} \\ K\ddot{R} &= \nabla^2 R - 4\pi Ne\ddot{\zeta} \end{aligned}\right\} \quad \dots \quad (24)$$

determine the problem

If a plane wave be propagated parallel to a line of force, P and ξ vanish, and by elimination of η and ζ between (23) and (24) we may obtain two equations which only contain P and Q For the sake of simplicity, we shall confine ourselves to the simple periodic motion.

Writing $-i\omega$ for d/dt and $-\omega^2$ for d^2/dt^2 and introducing symbols σ and Π defined by

$$\sigma_r = \eta + i\zeta, \quad \Pi_r = Q + iR,$$
$$\sigma_l = \eta - i\zeta, \quad \Pi_l = Q - iR,$$

we obtain from (23) and (24)

$$\left.\begin{aligned} \sigma_r\{(n^2 - \omega^2)\rho + He\omega\} &= e\Pi_r \\ \sigma_l\{(n^2 - \omega^2)\rho - He\omega\} &= e\Pi_l \end{aligned}\right\} \quad \dots\dots \quad (25),$$

$$\left.\begin{aligned} K\ddot{\Pi}_r &= \nabla^2\Pi_r - 4\pi Ne\ddot{\sigma}_r \\ K\ddot{\Pi}_l &= \nabla^2\Pi_l - 4\pi Ne\ddot{\sigma}_l \end{aligned}\right\} \quad \dots\dots \quad (26)$$

σ_r and σ_l may now be eliminated, and we derive thus from (25) and (26),

$$\left(K + \frac{4\pi Ne^2}{(n^2 - \omega^2)\rho + He\omega}\right)\frac{d^2\Pi_r}{dt^2} = \frac{d^2\Pi_r}{dx^2},$$

$$\left(K + \frac{4\pi Ne^2}{(n^2 - \omega^2)\rho - He\omega}\right)\frac{d^2\Pi_l}{dt^2} = \frac{d^2\Pi_l}{dx^2}$$

This gives for v_r the velocity of right-handed circularly polarized

* Maxwell, *Electricity and Magnetism*, Vol II p 227.

light, and for v_l the velocity of left-handed circularly polarized light,

$$\left.\begin{aligned}\frac{1}{v_r^2} &= K + \frac{4\pi Ne^2}{(n^2-\omega^2)\rho + He\omega}\\ \frac{1}{v_l^2} &= K + \frac{4\pi Ne^2}{(n^2-\omega^2)\rho - He\omega}\end{aligned}\right\} \quad . \quad . \quad . \quad (27),$$

and

$$\frac{1}{v_l^2} - \frac{1}{v_r^2} = \frac{8\pi e^3 NH\omega}{(n^2-\omega^2)^2\rho^2 - H^2e^2\omega^2} \quad . \quad (28).$$

In actual observation, it is difficult to apply a field much greater than 30,000 units Assuming for e/ρ the value $1 \cdot 6 \times 10^7$, we find for He/ρ, $4 \cdot 8 \times 10^{11}$ Also ω for green light is $1 \cdot 3 \times 10^{15}$. If n and ω differ by not less than the two-hundredth part of their period, so that $n - \omega = 6 \times 10^{12}$ or more, the second term in the denominator of (28) is equal to less than the six-hundredth part of the first and may be neglected

We may write under these circumstances, if v represents the velocity of light in the absence of a magnetic field,

$$\frac{1}{v_l^2} - \frac{1}{v_r^2} = \frac{2}{v}\left(\frac{1}{v_l} - \frac{1}{v_r}\right) = \frac{8\pi e^3 NH\omega}{(n^2-\omega^2)^2\rho^2}$$

The gyric coefficient (r) is equal to $\frac{1}{2}\omega\left(\frac{1}{v_r} - \frac{1}{v_l}\right)$ and hence

$$\begin{aligned} r &= -\frac{2\pi e^3 NH\omega^2 v}{(n^2-\omega^2)^2\rho^2}\\ &= -\frac{\beta\omega^2 vH}{(n^2-\omega^2)^2} = -\frac{\beta\omega^2 VH}{\mu(n^2-\omega^2)^2} \quad . \quad (29),\end{aligned}$$

where β is an appropriate coefficient, and μ the refractive index

If the free periods are much more rapid than those to which the observations apply, ω in the denominator of (29) may be neglected and r is approximately proportional to ω^2, which agrees with the experimental facts If the electron, the motion of which has been considered, is positive, the rotation has the opposite sign to H As we take a clockwise rotation as negative, this means that the rotation is right-handed when the light travels in the direction of a line of force (from North to South) If the vibrating electron is negative, the opposite is the case, and the turn of the plane of polarization is then in the same direction as that of the positive current in a solenoid having its lines of force coincident with that of the field

The right-hand side of (29) being inversely proportional to $(n^2-\omega^2)^2$ becomes abnormally great when the period of the transmitted light approaches the free period of the molecule, but the direction of rotation remains the same whether ω is greater or smaller than n.

When n is nearly equal to ω, equation (29) fails to be correct, and we must derive r from (28)

169. Connexion between the Zeeman effect and magneto-gyration. We might have derived the results of the last article directly from the general theory of refraction, taking account of the changes in the free periods due to the magnetic field. Equation (42) Chapter XI gives us for the velocity of a wave in a responsive medium·

$$\frac{1}{v^2} = K + \frac{4\pi e^2 N}{\rho(n^2 - \omega^2)} \quad . \quad . \quad . \quad . \qquad (30)$$

In the magnetic field, the free period $2\pi/n$ is altered, and is different for circular vibrations according as they are left-handed or right-handed. According to (19) of this Chapter we must therefore substitute $\rho n^2 \pm He\omega$ for ρn^2, the upper sign holding for the right-handed rotation. This introduced into (30) leads directly to (27)

The investigation of the last paragraph has been derived from an important paper by W Voigt, who first gave equations which are practically identical with, though in one respect more general than (28) Voigt adds a frictional term to the equations of motion, in order to include the phenomenon of absorption, but owing to the objections raised in Art 153 against the introduction of this term it has been omitted here

The importance of Voigt's work consists in the establishment of a simple and rational connexion between the Zeeman effect and magneto-gyric properties. Each free period of the molecule is divided by the magnetic field into two, one being dextro-gyric, and the other laevo-gyric. Each of these imposes a rotatory polarization in its own direction, the velocity of propagation being increased on the violet side and diminished on the red side. Consider a period on the red side of a Zeeman doublet. It is most affected by the least refrangible component, the effect being a *diminution* of velocity, hence the resulting photo-gyric effect is in the same direction as that of the *most* refrangible component. On the violet side the most refrangible component is the one that is most active, and as the effect is here an *increase* in velocity, it follows that the photo-gyric effect is also in this case in the direction of the most refrangible component of the Zeeman doublet. This is true for all vibrations which do not fall within the periods intermediate between those of the two components, where the effect is in the opposite sense as easily reasoned out in the same manner. Zeeman's observations on Sodium light show that the most refrangible component rotates in the direction of the solenoidal current, giving a magnetic force coincident with that of the field, and this is therefore the direction in which we should expect sodium vapour to

rotate the plane of polarization, except within a very narrow range close to the undisturbed period Observation confirms this.

We conclude our theoretical discussion by deducing a remarkable relation first brought forward on more speculative grounds by H. Becquerel. Not necessarily confining ourselves to single free periods, we may write equation (30)

$$\frac{1}{v^2}=K+\Sigma\frac{\beta}{(n^2-\omega^2)} \quad . \quad . \quad (30a),$$

where the summation extends to the different values of n for which β may also have different values.

With the ordinary notation for small quantities, we may put

$$\delta\frac{1}{v^2}=\frac{d}{dn}\left(\frac{1}{v^2}\right)\delta n,$$

or

$$\frac{2\delta v}{v^3}=-\frac{d}{dn^2}\left(\frac{1}{v^2}\right)\delta n^2.$$

If $1/v^2$ has the form of (30a) we may substitute differentiation with respect to ω^2 for differentiation with respect to $-n^2$ and hence

$$\frac{\delta v}{v^2}=\tfrac{1}{2}v\frac{d}{d\omega^2}\left(\frac{1}{v^2}\right)\delta n^2.$$

If δv represents the increase in the velocity of propagation of the laevo-gyric light due to the magnetic field, the gyric coefficient (r) is $\omega\delta v/v^2$ For δn^2 we may write, according to (19) and (20), $4\pi H\omega z$, so that

$$\begin{aligned} r &= \pi z H\omega v\frac{d}{d\omega}\left(\frac{1}{v^2}\right) \\ &= -2\pi\frac{zH\omega}{V}\frac{d\mu}{d\omega} \\ &= -2\pi\frac{zH\lambda}{V}\frac{d\mu}{d\lambda} \quad \ldots\ldots\ldots(31), \end{aligned}$$

where the refractive index μ has been substituted for V/v. This is Becquerel's equation*, which will be further discussed in the next article

170. Experimental Facts and their connexion with the theory. The magneto-gyric effects of the great majority of substances are in the positive direction, by which we mean that they are in the same direction as that of the solenoidal current producing the magnetic field If our theory is correct, this would mean that it is the negative electron which is the active vibrator, a result which we had already derived from the Zeeman effect The salts of iron form however a

* *C. R.* cxxv. p. 679 (1897).

notable exception, for it is found that those salts which are magnetic have a negative coefficient This at one time led to the belief that there might be a characteristic difference between dia-magnetic and para-magnetic bodies, the latter possessing a negative coefficient. The following table which has been given by H du Bois shows how far such a distinction is justified.

TABLE XV

Diamagnetic		Paramagnetic	
Dextro-gyric	Laevo-gyric	Dextro-gyric	Laevo-gyric
Potassium ferro-cyanide	Titanium chloride	Iron	Ferrous salts
Lead borate		Cobalt	Ferric salts
Water		Nickel	Potassium ferricyanide
Hydrogen		Oxygen	Chromium trioxide
The great majority of solid, liquid and gaseous substances		Nitric oxide	Potassium bichromate
		Cobalt salts	Potassium chromate
		Nickel salts	Cerium salts
		Manganese salts	Lanthanum salts
		Cupric salts	Didymium salts

It is notable that the three magnetic metals, iron, nickel and cobalt, have a positive gyric coefficient, which seems at first sight in direct contradiction to the suggested connexion But it has been found that for these metals $d\mu/d\lambda$ is positive, so that if Becquerel's law is generally true, the negative value of r might be explained. Titanium chloride is the only diamagnetic body which gives a negative r, but Titanium is a magnetic metal and therefore it is possible to argue that the dia-magnetism of chlorine overpowers the magnetism of Titanium, but that with regard to the gyric property the metal has the upper hand The same argument cannot however be used to explain the positive coefficient of oxygen, and the salts of cobalt, nickel and manganese. The subject is suggestive, but requires further experimental treatment. Should the negative coefficients be ultimately found to be confined to magnetic substances, it will not be necessary to assume that their vibrating electrons are positive The magnetic molecule may have a gyric property in virtue of its being magnetic, and the effects of this property would superpose themselves on the other effects, to which our theory has been confined. In the case of feebly magnetic substances, the Zeeman gyration may gain the upper hand, while in

chloride of Titanium, the pure magnetic vortex rotation may be superposed to be paramount. This might explain some of the discrepancies of the above Table. A theory of magnetic vortex-gyration has been given by Drude*

The following table given by H Becquerel† shows the magnitude of the rotation of different substances compared with carbon bisulphide, and gives also the values of the Zeeman coefficient calculated from (31).

TABLE XVI

Substance	Relative magneto-gyric coefficient	$\lambda \frac{d\mu}{d\lambda}$	$z \times 10^{-5}$
Oxygen	000146	$1\ 47 \times 10^{-5}$	5 98
Air	000159	$1\ 44 \times 10^{-5}$	6 64
Nitrogen	000161	$1\ 68 \times 10^{-5}$	5 74
Carbonic acid	000302	$2\ 00 \times 10^{-5}$	9 07
Nitrous oxide	000393	$4\ 85 \times 10^{-5}$	4 88
Water	308	$1\ 99 \times 10^{-2}$	9 33
Benzine	636	$4\ 88 \times 10^{-2}$	7 85
Phosphor trichloride	651	$4\ 71 \times 10^{-2}$	8 30
Carbon bisulphide	1 000	$9\ 71 \times 10^{-2}$	6 20
Liquid phosphorus	3 120	$2\ 52 \times 10^{-1}$	7 41
Titanium bichloride	−1 358	$9\ 96 \times 10^{-2}$	−2 16

Excluding the dextro-gyric titanium bichloride it will be noticed that the Zeeman coefficients for these substances, having widely different dispersions, are all of the same order of magnitude, thus giving a substantial confirmation of the correctness of Becquerel's law The numerical value of z in the above Table is about half that of the lowest and one quarter that of the highest number obtained directly from observations with luminous vapours as shown by Table XIV An apparent increase of electric mass, in the more complicated structures of molecules which do not give line spectra, is thereby suggested.

Absolute determinations of the magneto-gyric coefficients have been made for carbon bisulphide and for water. The most recent determinations, reduced to unit magnetic force, are, in minutes of arc

* *Lehrbuch der Optik*, p. 384

† *C R.* cxxv. p 683 (1897)

TABLE XVII

Bisulphide of Carbon

Lord Rayleigh*	for Sodium light,	$t=18°$,	$r=$	0′ 04200
Kopsel†	” ”	”		0′ 04199
”	” ”	$t=0°$,		0′ 04207
Becquerel‡	” ”	”		0′ 04341

Water

Arons§		$t=23°$,	0′ 01295

The gyric effect of thin films of iron, when magnetized to saturation, is enormous Its discoverer, Kundt, found it to be at the rate of a complete revolution for a thickness of 02 mm which gives 200,000° for one centimetre, an effect which is 290 million times greater than in bisulphide of carbon. Cobalt gives a value nearly as great as found by Du Bois, and nickel about half as great.

The magneto-gyric coefficient is in general roughly proportional to the square of the frequency, showing that the vibrations which chiefly determine it have a very small wave-length, but anomalous cases have been noted Kundt‖ observed that thin iron films rotate the plane of polarization of red light more than that of blue light, and Lobach¶ measured the rotational coefficients of iron, nickel and cobalt in different parts of the spectrum The diminution in the angle of rotation between $\lambda=6\ 7\times10^{-5}$ and $\lambda=4\ 3\times10^{-5}$ was found to be approximately for iron 45 %, for cobalt 23 %, and for nickel 41 %

An interesting confirmation of the theory given in Art 169 is obtained by the observation of the gyric effects in the neighbourhood of absorbing regions of the spectrum. As has been pointed out in that article, the introduction of the Zeeman effect into Sellmeyer's equation leads directly to the conclusion that on both sides of an absorption line, there is a strong magneto-gyric effect in the direction in which the more refrangible members of the Zeeman components rotate This fact had been observed in the neighbourhood of the sodium lines by Macaluso and Corbino** (the gyric coefficient being positive, *i e* the rotation in the direction of the current producing the field). It has been further extended and commented upon by H Becquerel††.

* *Collected Works*, Vol II p 360

† *Wied Ann* Vol XXVI p 456, 1885

‡ *Ann Chim Phys* Vol XXVII p 312, 1882

§ *Wied. Ann* Vol XXIV p 161, 1885

‖ *Wied Ann* Vol XXIII. p 237, 1884

¶ *Wied Ann* Vol XXXIX p 346, 1890

** *C R* CXXVII p 548, 1898

†† *C R* CXXVII p 647, 1898

171. Double Refraction at right angles to the lines of force. If a plane wave traverses the magnetic field at right angles to the lines of force, the vibrations parallel to the field are propagated with different velocities from those at right angles. This follows also from Sellmeyer's theory of refraction in combination with the Zeeman effect.

Taking the simplest case of a line split by the magnetic field into a Zeeman triplet, the outer components affect the velocity of light of the vibrations normal to the field, while the central component affects the vibrations parallel to the field. Approaching an absorption line from the less refrangible side, the first effect will be a diminution of the velocity of both components, but to a greater degree of that component which lies nearest, *i.e.* the vibration normal to the field. Similarly approaching the absorption line from the violet end, both components are accelerated, and it is again the component vibrating normally to the field which is most affected. Hence there is double refraction in such a sense that towards the red end the vibration parallel to the line of force is propagated most quickly and on the violet side the vibration normal to the field. This result was predicted by W. Voigt from the theory and verified experimentally by him in conjunction with Wiechert*.

To obtain an expression for the amount of double refraction to be expected, we write Sellmeyer's equation for the light vibrating normally to the field in the form

$$\frac{1}{v_n^2} = K + \frac{\beta'}{(n_r^2 - \omega^2)} + \frac{\beta'}{(n_l^2 - \omega^2)},$$

where

$$n_r^2 = n^2 + 4\pi z\omega H, \quad n_l^2 = n^2 - 4\pi z\omega H,$$

$$\frac{1}{v_n^2} = K + \frac{2\beta'(n^2 - \omega^2)}{(n^2 - \omega^2)^2 - 16\pi^2 z^2\omega^2 H^2}.$$

The vibrations parallel to the lines of force are undisturbed and hence

$$\frac{1}{v_p^2} = K + \frac{\beta}{n^2 - \omega^2}$$

For $H = 0$, the two expressions must agree, and hence $2\beta' = \beta$. Writing a^2 for

$$16\pi^2 z^2\omega^2 H^2/(n^2 - \omega^2),$$

we have

$$\frac{1}{v_n^2} - \frac{1}{v_p^2} = \frac{\beta}{n^2 - \omega^2 - a^2} - \frac{\beta}{n^2 - \omega^2}$$

$$= \frac{\beta a^2}{(n^2 - \omega^2)(n^2 - \omega^2 - a^2)}.$$

* *Wied. Ann.* Vol. LXVII. p. 345, 1899.

If we treat α^2 as a small quantity compared with $n^2 - \omega^2$ and reintroduce its value, we find

$$\frac{1}{v_n^2} - \frac{1}{v_p^2} = \frac{16\beta\pi^2 z^2 \omega^2 H^2}{(n^2 - \omega^2)^3}.$$

When the vibrations on the contrary are so near to the free undisturbed period that $n^2 - \omega^2$ is small compared with α^2

$$\frac{1}{v_n^2} - \frac{1}{v_p^2} = -\frac{\beta}{(n^2 - \omega^2)}$$

The double refraction is now in the opposite direction, and hence close to the free period the perpendicular vibrations are propagated more quickly when $n > \omega$, *i e.* on the side of lower frequency.

CHAPTER XIII.

TRANSMISSION OF ENERGY.

172. Propagation of Energy Energy may be transmitted through a surface either by the passage of matter in motion, or by the performance of work An example of the first kind of transference of energy is found in the propagation of heat When conduction of heat takes place in gases, the total quantity of energy carried by the molecules of matter traversing a surface in the direction in which the heat is propagated is greater than that traversing it in the reverse direction But we are not in this book concerned with this simple and direct method of carriage of energy

Waves propagated through elastic solids may carry energy across a surface when the motion of matter is tangential as, owing to the tangential forces, work will in general be done A transference of energy may result owing to the fact that though the velocities and stresses are alternately in one direction and the other, their product contains a part which is not periodic Such cases of transference of energy can only be accurately investigated when we have some clear idea as to the mechanism of the motion which is involved. This renders it necessary to study in the first instance some simple illustrations of work done in the transmission of waves through elastic bodies

173. Waves of pure compression or dilatation in a perfect fluid. We choose a perfect fluid in order to simplify the equations as much as possible Putting $n = 0$, and writing ξ, η, ζ for the displacements, equations (9) Art 132 become

$$\left.\begin{aligned} \frac{d^2\xi}{dt^2} &= v^2 \frac{d\delta}{dx} \\ \frac{d^2\eta}{dt^2} &= v^2 \frac{d\delta}{dy} \\ \frac{d^2\zeta}{dt^2} &= v^2 \frac{d\delta}{dz} \end{aligned}\right\} \quad \ldots \quad (1),$$

where v^2 is written for κ/ρ and

$$\delta = \frac{d\xi}{dx} + \frac{d\eta}{dy} + \frac{d\zeta}{dz} \quad . \qquad (2).$$

For the stresses we have according to (7) and (8) Art. 131

$$S = T = U = 0,$$

$$P = Q = R = v^2 D\delta,$$

where D now stands for the density

If we consider in the first place a plane wave, the displacements being parallel to the axis of x, the first of equations (1) becomes

$$\frac{d^2\xi}{dt^2} = v^2 \frac{d^2\xi}{dx^2},$$

which is satisfied by

$$\xi = A \sin \frac{2\pi}{\lambda}(x - vt)$$

If W represents the work done across the surface

$$\frac{dW}{dt} = -P\frac{d\xi}{dt}$$

$$= -v^2 D \frac{d\xi}{dx}\,\frac{d\xi}{dt},$$

where the negative sign had to be introduced on the right hand because P is taken as positive when it is a *tension* and acts therefore in the opposite direction to that in which the velocities are taken as positive

Substituting for ξ and confining our attention to the plane $x = 0$,

$$\frac{dW}{dt} = \frac{4\pi^2 v^3 D A^2}{\lambda^2} \cos \frac{2\pi}{\lambda} vt.$$

By integration

$$W = \frac{2\pi^2 v^3 D A^2}{\lambda^2}\left(t + \frac{\lambda}{4\pi v} \sin \frac{4\pi}{\lambda} vt\right)$$

The second term vanishes at intervals of time which are equal to half a complete period, and becomes more and more negligible as t increases Leaving this term out of account, we may write for the work transmitted through unit surface

$$W = \tfrac{1}{2} D V_1^2 vt \quad . \qquad . \ . \ (3),$$

where V_1 stands for the maximum velocity If the whole mass of air through which the waves have spread in time t had a velocity equal to the maximum velocity V_1, its kinetic energy would be equal to that transmitted through the surface As the average kinetic energy in a simply periodic wave is equal to half the maximum energy, only half the energy transmitted through the surface is in the kinetic form, the other half being potential

It is important to notice that the transmission of energy depends on the fact that the velocity and condensation (which is proportional to the pressure) are in the same phase The condensed portions of the fluid move in the direction in which the wave is propagated, and the rarefied portions in the opposite direction It is a consequence of this fact that the work done while the air moves forwards is not undone while the air moves backwards

We next take the case of waves diverging from a point. The motion to be considered belongs to an important class in which the velocities may be represented as the partial differential coefficients of the same function ϕ, called the velocity potential

Put $$\frac{d\xi}{dt} = \frac{d\phi}{dx}, \quad \frac{d\eta}{dt} = \frac{d\phi}{dy}; \quad \frac{d\zeta}{dt} = \frac{d\phi}{dz}$$

Equations (1) are now all contained in the simple equation

$$\frac{d^2\phi}{dt^2} = v^2\nabla^2\phi \qquad \ldots \quad (4)$$

For the stress P we have

$$\frac{dP}{dt} = v^2 D \frac{d\delta}{dt}$$
$$= v^2 D\nabla^2\phi$$
$$= D\frac{d^2\phi}{dt^2}$$

Hence by suitable choice of the constant of integration

$$P = D\frac{d\phi}{dt}.$$

If ϕ depends only on the distance r from a fixed point which acts as a source from which the vibrations emanate, we have

$$\frac{d\phi}{dx} = \frac{d\phi}{dr}\frac{x}{r},$$

$$\frac{d^2\phi}{dx^2} = \frac{d^2\phi}{dr^2}\frac{x^2}{r^2} - \frac{d\phi}{dr}\frac{1}{r} - \frac{d\phi}{dr}\frac{x^2}{r^3}$$

Changing similarly the variable in $\dfrac{d^2\phi}{dy^2}$, $\dfrac{d^2\phi}{dz^2}$, we find by addition

$$\nabla^2\phi = \frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr}$$
$$= \frac{1}{r}\frac{d^2 r\phi}{dr^2}$$

Equation (4) now becomes

$$\frac{d^2 r\phi}{dt^2} = v^2\frac{d^2 r\phi}{dr^2},$$

the solution of which is

$$r\phi = f(r - vt),$$

or confining ourselves to the simple periodic motion

$$\phi = \frac{A}{r} \sin \frac{2\pi}{\lambda}(r - vt)$$

This value of ϕ is therefore a solution of the differential equation (4). By differentiation

$$\frac{d\phi}{dr} = \frac{A}{r}\left\{\frac{2\pi}{\lambda}\cos\frac{2\pi}{\lambda}(r-vt) - \frac{1}{r}\sin\frac{2\pi}{\lambda}(r-vt)\right\} \qquad (5).$$

$\frac{d\phi}{dr}$ represents the velocity, which it appears does not vary inversely as the distance as might have been expected at first sight. At small distances from the origin the second term is the important one and the velocity varies inversely as the square of r. The origin itself is a singular point at which matter enters and leaves the space. The amount of matter passing per unit time through any sphere having the origin as centre is equal to $4\pi r^2 D d\phi/dr$ which, if r is small, is equal to $4\pi A D \sin \frac{2\pi}{\lambda} vt$, and this expression therefore represents the rate at which matter is introduced at the origin. At large distances the first term is the important one. If the difference of phase between any point very near the origin and one at a large distance r away from it, were calculated in the usual way and put equal to $2\pi r/\lambda$, we should commit an error equal to a quarter of a wave-length. This apparent change of phase of a right angle when points near a source and at some distance away from it are considered, has been already referred to several times (*e.g.* Art. 16). If we were to measure energy simply by the square of the amplitude, equation (4) would lead to the conclusion that the energy does not vary inversely as the square of the distance from the origin as is generally assumed. There is however no reason why it should vary according to the simple law, so long as the energy *transmitted* follows it. That this is actually the case may be proved as follows. The rate at which energy is transmitted through a sphere of radius r is

$$\begin{aligned}\frac{dW}{dt} &= 4\pi r^2 p \frac{d\phi}{dr} \\ &= 4\pi D r^2 \frac{d\phi}{dt} \cdot \frac{d\phi}{dr} \\ &= 2\pi D A^2 \sin(\omega t - lr)\left\{l \sin(\omega t - lr) - \frac{1}{r}\cos(\omega t - lr)\right\}.\end{aligned}$$

Integrating with respect to the time and leaving out periodic terms, we find

$$W = 2\pi\omega D A^2 l t$$

This expression does not contain r and hence the work transmitted through concentric spheres enclosing the origin is constant. It follows

that the work transmitted in a given time through unit surface varies inversely as the square of the distance. For a fuller treatment of the subject the reader is referred to Lord Rayleigh's treatise on Sound, Vol. II., Arts. 279 and 280.

174. Plane Waves of Distortion in an elastic medium. Let the displacements be parallel to the axis of z and be denoted by ζ, the wave normal being the axis of x. The only force which can do work across the plane xy is the tangential stress which in Art. 129 has been called T, and which according to (7) Art. 131 is equal to $nd\zeta/dx$, ξ being zero in the present case. The stress T has been taken to be positive when the portion of matter on the positive side of the plane yz acts on the matter which is on the negative side with a force directed along the positive axis of z. Hence for waves travelling in the positive direction, if W be the energy transmitted across unit surface,

$$\frac{dW}{dt} = -n\frac{d\zeta}{dx}\frac{d\zeta}{dt}.$$

If $$\zeta = A\sin\frac{2\pi}{\lambda}(x - vt)$$

and the coefficient of distortion n is replaced by v^2/D, we find, as in the case of the sound-wave, leaving out periodic terms,

$$W = \tfrac{1}{2}DV_1^2vt \quad \ldots \ldots \quad (6),$$

where V_1 denotes the maximum velocity.

175. Sphere performing torsional oscillations in an elastic medium. Consider displacements in an elastic medium defined by

$$\xi = 0, \quad \eta = -\frac{d\phi}{dz}, \quad \zeta = \frac{d\phi}{dy}$$

where

$$\phi = \frac{A}{r}\sin\frac{2\pi}{\lambda}(r - vt)$$

The displacements satisfy the condition

$$\frac{d\xi}{dx} + \frac{d\eta}{dy} + \frac{d\zeta}{dz} = 0,$$

which shows that there is no condensation.

Also
$$\nabla^2\eta = -\frac{d}{dz}\nabla^2\phi$$
$$= -\frac{1}{v^2}\frac{d}{dz}\frac{d^2\phi}{dt^2}$$
$$= \frac{1}{v^2}\frac{d^2\eta}{dt^2}.$$

Similarly
$$\nabla^2\zeta = \frac{1}{c^2}\frac{d^2\zeta}{dt^2},$$
so that equations (9), Art. 132, are satisfied if $c^2 = n/D$. The assumed system of displacements represents therefore a possible wave propagation, the waves being purely distortional.

As ϕ does not contain x, y, z explicitly,
$$\frac{d\phi}{dy} = \frac{d\phi}{dr}\cdot\frac{y}{r}, \quad \frac{d\phi}{dz} = \frac{d\phi}{dr}\cdot\frac{z}{r}.$$

Hence
$$x\xi + y\eta + z\zeta = -y\frac{d\phi}{dz} + z\frac{d\phi}{dy}$$
$$= 0.$$

It follows that the displacements at any point are at right angles to the radius vector drawn from the origin to that point. As there are no displacements parallel to the axis of x, the displacements are along circles drawn round OX as axis.

Let ρ be the distance of any point from the axis, so that $r^2 = x^2 + \rho^2$. We obtain the amount of the displacement by resolving η and ζ in a direction at right angles to ρ in a plane parallel to the plane of yz. This gives for the displacement:
$$\frac{-\eta z + \zeta y}{\rho} = \frac{1}{\rho}\left(z\frac{d\phi}{dz} + y\frac{d\phi}{dy}\right)$$
$$= \frac{z^2 + y^2}{\rho r}\frac{d\phi}{dr}$$
$$= \frac{\rho}{r}\frac{d\phi}{dr}.$$

The angular displacement obtained by dividing the actual displacement by ρ only depends on r, and is therefore the same at all points of a sphere having the origin as centre. Each such sphere performs torsional oscillations as if it were rigid. We may therefore imagine any one sphere to be actually rigid and the oscillations to be maintained by forces applied to this sphere. Our system of equations will then tell us how these oscillations are propagated outwards.

In the language of Optics the vibrations at any point are polarized in a plane passing through OX which is the axis of rotation. The angular displacements are
$$\frac{1}{r}\frac{d\phi}{dr} = \frac{2\pi A}{\lambda r^2}\cos\frac{2\pi}{\lambda}(r - ct) - \frac{A}{r^3}\sin\frac{2\pi}{\lambda}(r - ct) \quad \ldots\ldots \quad (7),$$
and are nearly equal to the first or second term of this expression respectively, according as r is very small or very large compared with $\lambda/2\pi$. Comparing large and small values of r, we have here the same change of phase of a right angle which has been noted in Art. 173.

The maximum angular displacement at a distance S from the origin as obtained from (7) is:

$$A\sqrt{4\pi^2S^2+\lambda^2}/\lambda S^3,$$

and this is the amplitude of oscillation which must be maintained at a sphere of radius S in order to cause an angular amplitude $2\pi A/\lambda r^2$ at a large distance. If the maintained angular amplitude is B, it follows that for large distances the angular amplitude is

$$\frac{BS^2}{r^2\sqrt{1+\lambda^2(4\pi^2S^2)^{-1}}}$$

The actual amplitude is obtained on multiplying this expression by $r\sin\theta$, where θ denotes the angle which r forms with OX. To calculate the energy communicated by the rigid sphere to the surrounding medium, we make use of the obvious proposition that the energy transmitted through all concentric spheres must be equal and we may therefore simplify the calculation by considering only a sphere of very large radius.

If we write $(V_1\sin\theta)/r$ for the maximum velocity at a large distance, the total energy transmitted through unit surface at any time is by (6)

$$W=\tfrac{1}{2}DV_1^2vt\sin^2\theta/r^2,$$

and the work transmitted through the complete sphere is

$$\int_0^\pi 2\pi Wr^2\sin\theta\,d\theta=\int_0^\pi \pi tDV_1^2v\sin^3\theta\,d\theta=\frac{4\pi}{3}DV_1^2vt \quad \ldots (8)$$

Substituting the value of V_1, we find for E, the total energy transmitted,

$$E=\frac{4\pi Dvt}{3\lambda}\,\frac{S^4}{1+\lambda^2(4\pi^2S^2)^{-1}}\left(\frac{2\pi Bv}{\lambda}\right)^2$$

$$=\frac{4\pi}{3}\,\frac{DSvt}{1+\lambda^2(4\pi^2S^2)^{-1}}\left(\frac{2\pi BSv}{\lambda}\right)^2$$

The bracket on the right-hand side represents the greatest velocity in the equatorial plane of the rigid sphere.

It should be noticed that the energy transmitted diminishes with increasing wave-length (*i.e.* increasing period) and this diminution is the more important the smaller the radius of the embedded sphere is compared with the wave-length.

176. Waves diverging from a sphere oscillating in an elastic medium. The problems discussed in this and the preceding article were first solved by W. Voigt*. Kirchhoff† considerably simplified the mathematical analysis and more recently Lord Kelvin‡

* *Crelle's Journal*, Vol. LXXXIX. p. 288.

† *Crelle's Journal*, Vol. XC. p. 34.

‡ *Phil. Mag.* Vol. XLVII. p. 480 and XLVIII. pp. 277, 388.

has treated the same question very completely, adding several new and interesting results. We imagine a sphere embedded in an elastic medium, to which it is rigidly attached, and performing periodic linear oscillations according to the formula $A \sin \omega t$ The reader is referred to Lord Kelvin's *Baltimore Lectures* for the complete solution of the problem, we shall here confine ourselves to the question of emission of energy in an incompressible medium For this purpose it is only necessary to consider the motion at a distance which is large compared with the radius of the sphere (S) and the wave-length. If the oscillations of the sphere take place along the axis of x, Kelvin's equations for the motion when r is very great, are

$$\left.\begin{aligned} \xi &= -\frac{3}{2} AS\left(\frac{x^2}{r^3} - \frac{1}{r}\right) \sin(\omega t - lr) \\ \eta &= -\frac{3}{2} AS\frac{xy}{r^3} \sin(\omega t - lr) \\ \zeta &= -\frac{3}{2} AS\frac{xz}{r^3} \sin(\omega t - lr) \end{aligned}\right\} \qquad \ldots \ldots \ldots (9).$$

These equations give:

$$\xi x + \eta y + \zeta z = 0,$$

showing that the vibrations take place at right angles to the radius vector. The symmetry of the expression for the displacements as regards y and z shows that the displacements take place in meridional planes. For the resultant oscillation we have

$$(\xi^2 + \eta^2 + \zeta^2)^{\frac{1}{2}} = \frac{3}{2r} AS \sin\theta \sin(\omega t - lr),$$

where θ denotes the angle between the radius vector and the axis of x. The sign of the square root which occurs on the right-hand side is determined by the fact that for $x = 0$, the last equation must agree with the first of the equations (9) We note that in this case there is not the change of phase of a right angle which occurs when the sphere performs torsional oscillations In the language of Optics, the sphere may be said to send out polarized light, the vibrations being in meridional planes and at right angles to the ray The amplitude is a maximum in the equatorial plane, zero along the axis, and varies in intermediate positions as $\sin\theta$ If as in the last article, we write $V_1 \sin\theta / r$ for the maximum velocity, we may apply (8) directly to obtain the transmitted energy which is

$$\begin{aligned} E &= \frac{4\pi}{3} D V_1^2 vt \\ &= 3\pi D\omega^2 A^2 S'^2 vt. \end{aligned}$$

The emission of energy is therefore inversely proportional to the square of the period

177. Divergent Waves of Sound. The theory of Sound furnishes several important applications of the communication of energy from a vibrating body to a surrounding medium. If a stretched string vibrates backwards and forwards, the air which is compressed on one side is able to flow round the string and to diminish the rarefaction which tends to form behind the string. Under these circumstances comparatively little energy escapes in the form of sound-waves. Stokes* calculated the emission of the actual sound and compared it with that which would have been emitted if the lateral motion in the neighbourhood of the string were omitted. For a piano string of .02 inch radius sounding the middle C (wave-length about 25 inches) it appears that the prevention of the lateral motion would increase the intensity 40,000 times. This, as Stokes points out, shows the importance of sounding-boards, the broad surface of which is able to excite intense vibrations even though the motion itself is small. The following experiment may be described in Stokes' own words. "The increase of sound produced by the stoppage of lateral motion may be prettily exhibited by a very simple experiment. Take a tuning-fork, and holding it in the fingers after it has been made to vibrate, place a sheet of paper, or the blade of a broad knife, with its edge parallel to the axis of the fork, and as near to the fork as conveniently may be without touching. If the plane of the obstacle coincide with either of the planes of symmetry of the fork, as represented in section at *A* or *B*, no effect is produced, but if it be placed in an intermediate position, such as *C*, the sound becomes much stronger."

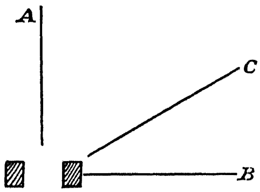

Fig. 177.

The motion of air round the sounding body is the more effective the shorter the wave-length. Were the length of the wave infinitely great, the air would move like an incompressible fluid backwards and forwards round the source of sound, and there would be no emission of energy once this motion is established. Stokes shows by applying the analysis to the case of vibrating spheres, that this is the explanation of an experiment due to Leslie, in which the sound of a bell placed in a partially exhausted receiver is diminished by the introduction of hydrogen.

178. Scattering of Light by small particles. In Arts. 175 and 176 the disturbance in a medium has been calculated on the supposition that at some given surface the *motion* is prescribed. There

* *Phil. Trans.* Vol. CLVIII. p. 447 (1868), Rayleigh, *Theory of Sound*, Vol. II. p. 306.

is a corresponding problem treated by Lord Rayleigh* in which within a given space the forces are prescribed Rayleigh finds that if a force $DZ \sin \omega t\,dx\,dy\,dz$ acts on an element of volume $dx\,dy\,dz$ having density D, the resulting disturbance may be expressed as

$$\zeta' = \frac{TZ \sin\theta}{4\pi v^2} \frac{\sin(\omega t - lr)}{r} \tag{10}$$

In this equation ζ' measures the displacement which lies in the plane passing through r and z and at right angles to r, and θ is the angle between z and r

The effect of a small suspended particle on a wave of light passing over it, was found by Lord Rayleigh with the help of the above result. The simplest supposition to make is that the difference in the optical property of the particle is due to a difference in density only. Let the primary vibration be

$$\zeta = \sin(\omega t - lx),$$

and the disturbing particle be at the origin The acceleration at that place is $-\omega^2 \sin \omega t$, and the corresponding force in the undisturbed medium $-D\omega^2 \sin \omega t$, to maintain the same vibration if the density is D', would require a force $-D'\omega^2 \sin \omega t$ The difference between these quantities or $-(D'-D)\,\omega^2 \sin \omega t$ measures the additional force which must be applied in order that the wave should pass undisturbed The actual effect is that of the undisturbed wave with the addition of that produced by a disturbing force equal and opposite to that which would annul the disturbance Substituting in (10) we find for the disturbance

$$\zeta' = \frac{D'-D}{D} \cdot \frac{\pi T \sin\theta}{\lambda^2 r} \sin(\omega t - lr)$$

The result shows that the scattered light is polarized, the plane of polarization being the plane containing the primary ray and the scattered ray which we shall call the principal plane The intensity of the scattered light varies as $\sin^2\theta$

If the direction of vibration in the primary ray is at right angles to the principal plane, we must put $\sin\theta = 1$, and if the direction of vibration is in the principal plane, $\sin\theta = \cos\phi$, where ϕ is the angle between the primary and secondary rays

If the incident light is unpolarized, the result will be the same as if there were two overlapping beams, one vibrating in the principal plane and the other at right angles to it Hence if the total intensity of the incident light be unity, so that each of the two overlapping beams has an intensity 1/2, the intensity of light in any direction is given by

$$\frac{(D'-D)^2}{D^2}(1+\cos^2\phi)\frac{\pi^2 T^2}{2\lambda^4 r^2} \tag{11}.$$

* *Collected Works*, Vol III p 166

If there is more than one particle, they will act independently of each other, owing to the constant change in their relative positions We need only therefore multiply the above expression by the total number of particles The explanation of the blue colour of the sky is contained in (11), for it shows that the intensity of the scattered light is much stronger for blue than for red light The observed polarization of the sky also generally corresponds to that indicated by the theory The minor differences which have been noticed are no doubt due to the fact that each part of the sky receives light not only directly from the sun, but also by reflexion from the illuminated portions of the earth and from the atmosphere itself Lord Rayleigh has also treated the same problem from the point of view of the electromagnetic theory, in which the differences in the electric properties of the suspended particles have to be considered If the differences between the dielectric constant of the particle and the medium are small, the above equations remain correct if D' and D denote the dielectric constants In the case of spheres, Lord Rayleigh solves the problem independently of the smallness of the difference in electric properties and finds that if K' and K denote the dielectric constants, we must, in the expression (11), replace $(D' - D)/D$ by

$$3(K' - K)/(K' + 2K)$$

A question of considerable importance may be raised in connexion with the nature of the bodies which scatter the light It was believed at one time that small particles of dust were responsible for the blue colour of the sky, but Lord Rayleigh* gave good reason to believe that the molecules of air themselves are sufficient for the purpose We may briefly indicate the argument by means of which the necessary data for the discussion of the problem are obtained The whole emission of energy by the small disturbed region T may be calculated by the method we have already employed in Arts 174 and 175 and is found to be

$$\frac{8\pi^3}{3} q^2 \frac{T^2}{\lambda^4} E,$$

where E is the energy transmitted through unit surface in the primary wave, and q measures a quantity which in the elastic solid theory of light is $(D' - D)/D$. If there be N similar particles per unit volume, a layer of thickness dx and unit area contains Ndx particles. Hence for dE, the diminution of the energy in the primary wave consequent on the scattering of light in the layer dx, we have

$$\frac{1}{E}\frac{dE}{dx} = -\frac{8\pi^3 N}{3} q^2 \frac{T^2}{\lambda^4}.$$

* *Collected Works*, Vol. IV p 397

Hence
$$E = E_0 e^{-kx}$$
where
$$k = \frac{8\pi^3 N}{3} q^2 \frac{T^2}{\lambda^4} \quad \ldots \quad (12).$$

The quantity k may be obtained by observation

We also reproduce Lord Rayleigh's investigation concerning the effect of the particles on the phase of the transmitted light Let the incident light be represented by $\cos(\omega t - lr)$ and consider a stratum at a distance x from the point O at which the light is to be estimated. Let A be the pole of O in the stratum and P the position of a scattering particle If ρ be the distance AP, the number of particles in the annular space between circles of radius ρ and $\rho + d\rho$ is $2\pi\rho N d\rho dx$. Also $r^2 = x^2 + \rho^2$ so that $\rho d\rho = r dr$ All these particles affect the vibration at O equally so that the amplitude caused by them is

$$N dx \frac{\pi q T}{r\lambda^2} \cos(\omega t - lr)\, 2\pi r dr.$$

Integrating this over the whole stratum, the limits of r being x and ∞, we find for the resultant amplitude

$$N dx \frac{\pi q T}{\lambda} \sin(\omega t - lr).$$

The primary wave may be supposed to advance undisturbed, so that the disturbance at O is

$$\cos(\omega t - lr) + NT dx \frac{\pi q}{\lambda} \sin(\omega t - lr)$$

or
$$\cos(\omega t - lr - \delta)$$
where neglecting the square of δ
$$\delta = NT dx \frac{\pi q}{\lambda}.$$

If μ be the refractive index of the medium as modified by the particles, $(\mu - 1)\,dx$ measures the change in optical length which is equal to $\lambda\delta/2\pi$ so that

$$(\mu - 1) = NTq/2 \quad \ldots \quad (13).$$

Introducing this expression into (12) we find

$$k = \frac{32\pi^3 (\mu - 1)^2}{3N\lambda^4} \quad (14),$$

which gives a relation in which the *number* of scattering particles is the only quantity which is not obtained by experiment or observation. There is some uncertainty at present in the value for k to be substituted in (14) Rayleigh founds his arguments on Bouguer's estimate giving ·8 for the fraction of light transmitted through the atmosphere from a star in the zenith Taking for N, Maxwell's estimate of 19×10^{18}, it appears that the actual transmission through the atmosphere is only about three

times less than that calculated from molecular diffraction without any allowance for the scattering by larger particles of dust Rayleigh concludes that "the light scattered from the molecules would suffice to give us a blue sky, not so very greatly darker than that actually enjoyed" If N be regarded as unknown we may use (14) to give a lower limit to its value, and it is thus found that N cannot be greatly less than the above numbers 19×10^{18} But N is certainly greater than was estimated by Maxwell The subject is very fully discussed by Lord Kelvin*, who concludes from the fact that a grayish haze is seen even on a clear day "that the want of perfect clearness of the lower regions of our atmosphere is in the main due to suspended particles too large to allow approximate fulfilment of Rayleigh's law of fourth power of wave-length" Kelvin considers it probable that not so much as a quarter or a fifth of Bouguer's degree of opacity is due to the ultimate molecules of air. We must refer to the Baltimore Lectures for the discussion by Lord Kelvin of Mr Majorana's observations on the luminosity of the sky on the summit of Mount Etna and Sella's observations on Monte Rosa Calling f the proportion of light due to the ultimate molecules of air in the sky over Mount Etna, and f' the same proportion in the sky over Monte Rosa, as deduced from the measurements of Majorana and Sella, Lord Kelvin sums up his discussion in the equations .

$$N = \frac{3.12}{f} \cdot 10^{19} = \frac{5.97}{f'} \, 10^{19} \quad . \qquad (15)$$

It would only be if the whole light of the sky were due to molecular diffraction that f or f' could have so great a value as unity. This argument may be taken to support Lord Kelvin's own estimate of $N = 10^{20}$ as the lower limit for the number of molecules in one cubic centimetre under normal conditions.

But this is probably too high an estimate. If we assume the charge of an electron to be known, the number N may be calculated at once In the electrolysis of water, one gram of hydrogen is set free by the passage of $\cdot 959 \times 10^4$ electromagnetic units of electricity Hence if N be the number of molecules, or $2N$ the number of atoms, in one cubic centimetre, and ρ the density under normal conditions

$$2Ne = \cdot 959 \times 10^4 \rho.$$

Substituting $\qquad \rho = 8.95 \times 10^{-5}$

we obtain $\qquad Ne = \cdot 429.$

The best estimate of e we possess is that of H A. Wilson† obtained by a modification of J. J. Thomson's method According to Wilson

* *Baltimore Lectures*, p 301 seq

† *Phil Mag* v p 429 (1903)

the quantity of electricity carried by the electron is very nearly 10^{-20} electromagnetic units This gives

$$N = 4\,3 \times 10^{19}.$$

This number agrees very well with (15) provided f and f' are unity. We are justified in concluding that at elevations of 14,000 feet and at zenith distances not greater than 60° (30° being the lowest elevation of the sun at which Majorana's observations were carried out) the light of the sky is almost entirely due to scattering by the ultimate molecules of the air It should be noticed that the method of calculating the retardation caused by the scattering particles which has led to (13) fails to give an account of absorption To explain the latter we should require a second approximation and take account of a difference in phase between the primary and secondary rays

179 Transmission of energy by electromagnetic waves. When we consider the transmission of energy in an electromagnetic field we are met at once by the difficulty that we are ignorant of the mechanism by which electromagnetic action is propagated Hence we cannot obtain an expression for the work done across a surface, unless we specify by means of an hypothesis what are the actual displacements of a medium subject to magnetic or electric forces In the special cases we have to consider here, the difficulty may be turned According to Maxwell a medium of permeability μ, subject to a magnetic force which at any place is H, possesses energy which per unit volume is measured by $\mu H^2/8\pi$. Similarly if K be the dielectric constant and E the electric force, $KE^2/8\pi$ is the electric energy also per unit volume. Consider now a plane wave propagated with velocity v in the direction of the axis of z

Equations (22) Art 137, which determine P, Q, the components of electric force in terms of α, β, the components of magnetic force, give us

$$KE^2 = K(P^2 + Q^2) = K\mu^2 v^2 (\alpha^2 + \beta^2)$$

As $K\mu = 1/v^2$ and $\alpha^2 + \beta^2$ measures the square of magnetic force, it follows that

$$KE^2 = \mu H^2,$$

so that we may, in the case considered, write for the energy per unit volume either $\mu H^2/4\pi$ or $KE^2/4\pi$. The wave need not be homogeneous and may be either plane or elliptically polarized or not polarized at all. Consider now a wave-front to the right of which the medium is at rest, to coincide with the plane of xy at the time $t = 0$ After time t, the wave-front will be at a distance vt from the origin, and the energy which has crossed unit surface of the plane of xy will be that contained in the volume having the unit surface as base and as length the distance vt measured along the axis of z If the magnetic force to the left of the wave-front is of the form

$$H = H_0 \cos(\omega t - lz),$$

the average value of H^2 is equal to $\frac{1}{2}H_0^2$ Hence the energy which has crossed unit surface in time t, putting $\mu = 1$, is $H_0^2 vt/8\pi$. The work done across a small surface of any wave-front cannot depend on the question whether the wave is plane or not We are therefore justified in using the expression obtained whenever the electromagnetic disturbance follows the simple periodic law

We next treat of a simple case in which we can trace the loss of energy of a radiating source We adapt for this purpose the results of Art 175, substituting the magnetic force for the displacement, so that we may write

$$\alpha = 0, \quad \beta = -\frac{d\phi}{dz}, \quad \gamma = \frac{d\phi}{dy} \qquad (16),$$

where

$$\phi = \frac{A}{r} \sin (lr - \omega t) \qquad (17)$$

These equations satisfy (19) of Art 135, and represent therefore a possible distribution of magnetic force It follows from the results of Art. 175 that the magnetic force at a distance ρ from the axis of x is $\rho \dfrac{d\phi}{dr} \Big/ r$ and that the lines of magnetic force are circles having OX as axis

We next consider a region round the origin so small that the phase at all points lying in it may be considered identical with that at the origin Within this region we may write

$$\phi = -\frac{A}{r} \sin \omega t$$

The components of current are obtained from (12) Art. 135

$$4\pi u = \frac{d\gamma}{dy} - \frac{d\beta}{dz} = \frac{d^2\phi}{dy^2} + \frac{d^2\phi}{dz^2}$$

$$= -\frac{d^2\phi}{dx^2},$$

$$4\pi v = -\frac{d^2\phi}{dydx},$$

$$4\pi w = -\frac{d^2\phi}{dzdx}$$

For the electric forces as obtained from 17, Art 134, we have

$$KP = -\frac{d}{dx}\frac{d^{-1}}{dt}\frac{d\phi}{dx} = \frac{d}{dx}\frac{Ax}{\omega r^3}\cos \omega t,$$

$$KQ = -\frac{d}{dy}\frac{d^{-1}}{dt}\frac{d\phi}{dx} = \frac{d}{dy}\frac{Ax}{\omega r^3}\cos \omega t,$$

$$KR = -\frac{d}{dz}\cdot\frac{d^{-1}}{dt}\frac{d\phi}{dx} = \frac{d}{dz}\frac{Ax}{\omega r^3}\cos \omega t$$

The electric forces close to the origin are therefore derivable from a potential

$$-\frac{Ax}{\omega r^3}\cos\omega t$$

If a quantity of electricity $-e$ is placed on the axis of x at a distance $\frac{1}{2}h$ from the origin, and similarly a quantity $+e$ at the same distance on the negative side, the electrostatic potential of such a so-called doublet is known to be $-ehx/r^3$. Hence we may represent the electric forces in the case we are considering by means of a doublet if we make

$$eh = (A\cos\omega t)/\omega$$

The system of waves represented by (16) and (17) may therefore be considered to be produced by a vibrating electric doublet at the origin, the two charges oscillating in the same period, reaching a maximum distance h and crossing at the origin. The quantity eh is called the moment of the doublet, for which we may write $M\cos\omega t$.

The total energy dissipated per unit time by the vibrations of such a doublet may now be calculated. At a large distance the magnetic forces are

$$\beta = \frac{lz}{r^2}M\omega\cos(lr-\omega t),$$

$$\gamma = -\frac{ly}{r^2}M\omega\cos(lr-\omega t),$$

from which we obtain,

$$\beta^2+\gamma^2 = l^2\omega^2M^2\sin^2\theta\cos^2(lr-\omega t)/r^2;$$

where θ is the angle between the radius vector and the axis of x. Hence through unit surface of large spheres the amount of energy which passes in time t is $l^2\omega^2M^2\sin^2\theta\,\omega t/8\pi r^2$. Integrating this over the whole sphere, this becomes $\frac{1}{3}l^2\omega^2M^2vt$, or expressing l and ω in terms of τ and v (where τ represents the time of vibration), we finally obtain for the energy sent out in time t by the vibrating doublet: $16\pi^4M^2t/3v\tau^4$.

In order to form some numerical estimate of this loss of energy consider the positive charge to remain fixed at the origin and the negative charge to vibrate according to the law $h\cos\omega t$. The maximum energy of the negative electron is $\frac{1}{2}m\omega^2h^2$ or $\frac{1}{2}m\omega^2M^2/e^2$, if m denotes its mass. From this we calculate the fraction of the maximum energy which is lost in a complete vibration taking up a time $2\pi/\omega$, to be $8\pi^2e^2/3\lambda m$. The ratio e/m is approximately known to be 10^7 and for e we may substitute 10^{-20}. This gives the loss of energy as being: $2.7\times10^{-12}/\lambda$. For violet light we have $\lambda = 4\times10^{-5}$, so that in each period a particle sending out such light would lose less than the millionth part of its energy. The motion of the particle, taking

account of the loss of energy by radiation, would have to be represented by the expression $he^{-\kappa t}\cos\omega t$, where the coefficient κ may be calculated from the data obtained. At each vibration there is a fractional diminution of the maximum velocity equal to $\kappa\tau$ and a fractional diminution of the energy equal to $2\kappa\tau$. Hence $1/\kappa\tau$ which is the number of vibrations in which the amplitude diminishes in the ratio $1 : e$ is $10^7/14$ or approximately 700,000 The small diminution in vibratory energy consequent on radiation justifies the criticisms made in Art. 151 and 153 respecting the introduction of a frictional term in the equations accounting for the so-called anomalous dispersion.

Consider now two similar doublets with their axes at right angles to each other, the positive electron being stationary and the negative oscillating in one case according to the law $h\cos\omega t$ and in the other according to the law $h\sin\omega t$. The electromagnetic effect will be the same as that of a single electron revolving with uniform speed in a circle of radius h, the loss of energy for each of the two vibrations at right angles to each other is that given above, and hence the total loss per unit time of such an electron revolving in time τ is $32\pi^4 M^2/3v\tau^4$ We arrive therefore at the remarkable conclusion that an electric charge describing a circle with uniform speed radiates energy, and as the speed is constant, the radiation can only depend on the acceleration which is directed to the centre Writing f for the acceleration we have

$$f = \omega^2 h = \frac{4\pi^2 h}{\tau^2},$$

$$\therefore\ f^2 = \frac{16\pi^4 h^2}{\tau^4} = \frac{16\pi^4 M^2}{e^2\tau^4}.$$

We obtain therefore $\frac{2}{3}e^2f^2/v$ for the loss of energy in unit time

This expression which is here proved for the special case that the acceleration is at right angles to the motion holds generally so long as the velocity is small compared with the velocity of light For a more detailed discussion, the reader is referred to Larmor's *Aether and Matter*, Chapter XIV.

180. Group Velocity. The preceding results have been deduced under the supposition that the medium propagates waves of different lengths with equal velocities. In such media, plane waves travel without change of type and the energy must be propagated with the same velocity as the wave. There are however important cases in which the supposition of equal wave velocities does not hold, and these bring out some important features in the mechanism of wave propagation If we watch a group of waves travelling over a sheet of water, we notice that the group as a whole does not move forward

so rapidly as the individual waves These seem to approach the front of the group and die out as they pass through it The explanation of this fact is obtained by considering the different rate at which water waves are propagated according to their length, but taking it simply as a fact, it shows that the energy of the waves can only travel on the average as quickly as the group, and must therefore travel more slowly than the waves

A group of waves always necessarily involves the superposition of waves of different lengths, because we can only identify any particular portion of the group by its distinction from the rest, either by a difference in the distance from crest to crest of two successive waves or by a difference in amplitude The analytical representation in both cases involves waves of different lengths.

Let two trains of waves be represented by $\cos l(Vt-x)$ and $\cos l'(V't-x)$ and their resultant by

$$\cos l(Vt-x)+\cos l'(V't-x)$$

which is equal to

$$2\cos\left\{\frac{l'V'-lV}{2}t-\frac{l'-l}{2}x\right\}\cos\left\{\frac{l'V'+lV}{2}t-\frac{l'+l}{2}x\right\} \qquad (18)$$

The first factor passes through its period in a time $4\pi/(l'V'-lV)$ while the periodic time of the second is $4\pi/(l'V'+lV)$ If $l'-l$ and $V'-V$ be sufficiently small, it takes many periods of the second factor to produce an appreciable difference in the first factor Hence we may say that the resultant effect is that of a wave having a length approximately equal to that of either train of waves, and an amplitude which varies slowly The velocity of the waves in this group is $(l'V'+lV)/(l'+l)$ or to the first approximation, equal to that which corresponds to a wave-length $2\pi/l$. To find the velocity of the group we must fix our mind on some special feature which may be chosen to be the maximum amplitude. For $t=0$, this lies at the origin, and generally the amplitude has its maximum whenever

$$(l'V'-lV)t-(l'-l)x=0 \qquad \ldots\ldots (19).$$

The highest point of the wave travels forward therefore at a rate which is $(l'V'-lV)/l'-l$ or, on the supposition of nearly equal values of l and l', we may write for the group velocity

$$U=\frac{dlV}{dl}=\frac{dV/\lambda}{d1/\lambda}$$

$$=V-\lambda\frac{dV}{d\lambda} \qquad \ldots\ldots (20),$$

showing the dependence of the group velocity on the variation of the wave velocity with the wave-length.

The explanation of the propagation of groups was first given by Stokes while Osborne Reynolds pointed out its connexion with the propagation of energy. The more general equation (20) is due to Lord Rayleigh* Professor Lamb† has given an instructive proof of this equation, in which one part of the group is distinguished from the rest, not by a difference in amplitude, but by a difference in wave-length Let the group consist of waves approximately of the simply periodic character but with a gradual change in the distance from crest to crest The group velocity will be the velocity with which a particular distance between two successive crests moves The wave-length λ may here be considered to be a function of x and of t The rate of change of λ at a point which moves with velocity $\frac{dx}{dt}$ is by the rules of the differential calculus $\frac{\partial\lambda}{\partial t}+\frac{\partial\lambda}{\partial x}\frac{dx}{dt}$ If the velocity of the point is equal to the group velocity, then by the definition of U the wave-length is constant, hence

$$\frac{\partial\lambda}{\partial t}+U\frac{\partial\lambda}{\partial x}=0 \qquad (21)$$

Now let the point move with velocity V, i e follow one crest The next crest will move with velocity $V+\lambda\frac{\partial V}{\partial x}$ or $V+\lambda\frac{\partial V}{\partial\lambda}\cdot\frac{\partial\lambda}{\partial x}$, hence $\lambda\frac{dV}{d\lambda}\cdot\frac{d\lambda}{dx}$ measures the rate at which the wave-length increases. This gives

$$\frac{\partial\lambda}{\partial t}+V\frac{\partial\lambda}{\partial x}=\lambda\frac{\partial V}{\partial\lambda}\frac{\partial\lambda}{\partial x} \quad . \quad . \qquad (22).$$

By combining (21) and (22), we return to equation (20)

One word of caution may be necessary, when we speak of the velocity of a group, we do not mean to imply that the whole of the group moves forward without any alteration just as if it were a single wave. The first variable factor of (18) always has its maximum value when the condition (19) is satisfied, but the second factor continuously changes and at intervals of time which are equal to half a period, this factor is alternatively ± 1, so that the maximum amplitude after such an interval is converted into a minimum The essential point is, that at periodically recurring intervals, the group regains its original feature, and the distance through which the group has moved forward divided by the interval is called the velocity of the group

* *Collected Papers*, Vol I p 322, *Sound*, Vol I § 191 and Appendix.

† *Proc. London Math. Soc*, Sec. II Vol I. p 473 (1904).

The following Table given by Rayleigh is interesting as giving the relation between group and wave velocities in particular cases

$V\propto\lambda$,	$U=0$	Reynolds' disconnected pendulums.
$V\propto\lambda^{\frac{1}{2}}$,	$U=\frac{1}{2}V$	Deep water gravity waves
$V\propto\lambda^{\circ}$,	$U=V$.	Aerial waves etc
$V\propto\lambda^{-\frac{1}{2}}$,	$U=\frac{3}{2}V$.	Capillary water waves
$V\propto\lambda^{-1}$,	$U=2V$.	Flexural waves in elastic rods or plates

The last two examples show that it is possible for the group to travel more quickly than the individual wave

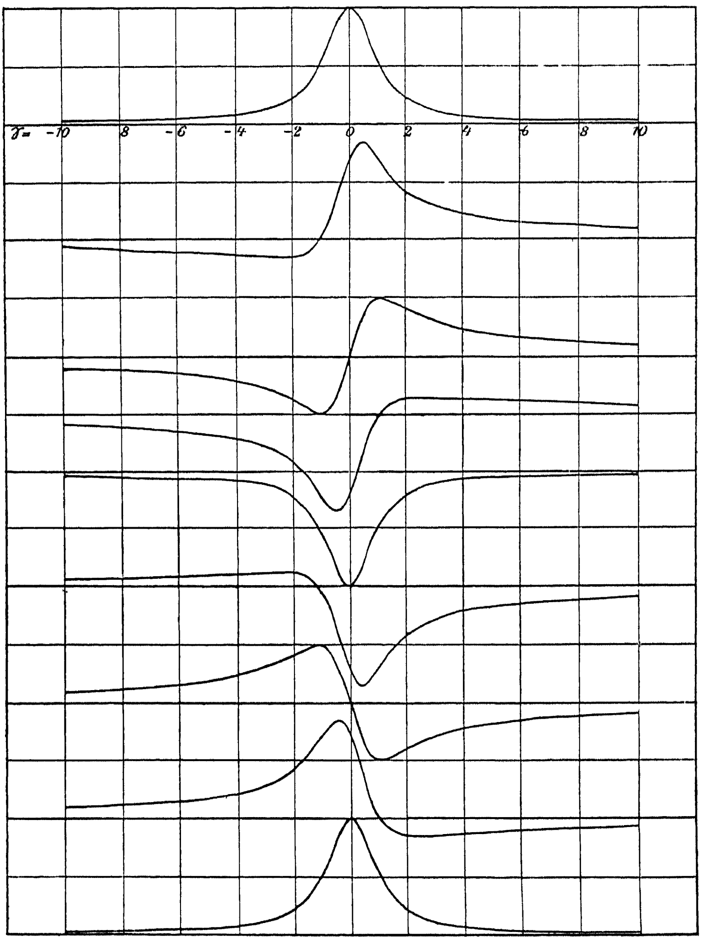

Fig. 178.

When the law connecting V and λ is $V = a + b\lambda$, the group velocity is independent of λ and the variations in the shape of the group may in special cases be followed out in detail* Fig 178 represents the successive stages of a group, the shape of which is represented by the equation

$$y = \frac{h^2}{h^2 + x^2}$$

An interesting question arises in the case of the propagation of light within an absorption band As explained in Art. 154 the wave velocity may increase with diminishing wave-length. In that case let V_0 and λ_0 represent the velocity and wave-length in vacuo, and let $dV/d\lambda_0$ be negative As $V\lambda_0 = V_0\lambda$, we obtain by differentiation with respect to V

$$\lambda_0 + V\frac{d\lambda_0}{dV} = V_0\frac{d\lambda}{dV},$$

or

$$\lambda + \frac{V^2}{V_0}\frac{d\lambda_0}{dV} = V\frac{d\lambda}{dV}.$$

As the second term on the left-hand side is negative, it follows that $\lambda\dfrac{dV}{d\lambda} > \lambda$, which shows that the group velocity is in the opposite direction to the wave velocity. If there is a convection of energy forward, the waves must therefore move backwards In all optical media where the direction of the dispersion is reversed, there is a very powerful absorption, so that only thicknesses of the absorbing medium can be used which are smaller than a wave-length of light. Under these circumstances it is doubtful how far the above results have any application But Professor Lamb† has devised mechanical ar-

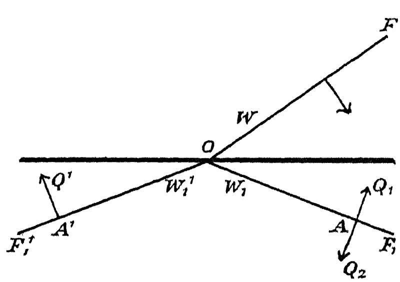

Fig 179

rangements in which without absorption there is a negative wave velocity One curious result follows the deviation of the wave on

* Schuster, *Boltzmann, Festschrift* p 569

† *Proceedings London Math Soc* Sec. II. Vol. I. p 473 (1904).

entering such a medium is greater than the angle of incidence, so that the wave normal is bent over to the other side of the normal as indicated in Fig 179 This is seen at once by considering that the traces on the refracting surface of WF and W_1F_1, the incident and refracted wave-fronts, must move together If we were to draw the wave-front in the usual way parallel to $W_1'F_1'$ and the waves moved backwards in the direction $A'Q'$, the intersection O of the refracted wave and surface would move to the left, while the intersection of the incident wave moved to the right. By drawing the refracted wave-front in the direction W_1F_1 the required condition can be secured. The individual waves move in the direction AQ_1 but the group moves in the direction AQ_2.

CHAPTER XIV

THE NATURE OF LIGHT

181. Application of Fourier's theorem. Gouy's treatment. We shall discuss in this chapter the conclusions that have been drawn with respect to the nature of the vibratory motion which leaves each molecule of a luminous body The difficulty we meet in interpreting the results of observation is partly due to the fact that the luminous disturbance which reaches us is the result of the superposition of the disturbances coming from a great number of irregularly distributed molecules, but the changes which optical appliances, used for purposes of observation, impress on the disturbance, have also formed a not infrequent source of confusion The light leaving an instrument such as a spectroscope must not be supposed to be identical in character with that entering it

The mathematical investigation is in many cases simplified by an application of Fourier's theorem. We consider a ray of plane polarized light and fix our attention on a point P over which the disturbance passes If the velocity at P be v, we may, in the most general case, express it as a function of the time, $f(t)$ Let us follow the motion from a time $t=0$, to a time $t=T$. According to Fourier's theorem, which has already been explained in Art 10, we may write

$$f(t) = a_0 + a_1 \cos(2\pi t/T) + a_2 \cos(4\pi t/T) + a_3 \cos(6\pi t/T)$$
$$+ b_1 \sin(2\pi t/T) + b_2 \sin(4\pi t/T) + b_3 \sin(6\pi t/T) \quad \ldots \quad (1)$$

Assuming that it is always possible to express v in terms of such a series, we may easily determine the value of any coefficient a_s by multiplying both sides by $\cos(2\pi st/T)$ and integrating between the limits 0 and T. It will be found that on the right-hand side all integrals have the same value at both limits except that one which has a_s for coefficient.

We find similarly any coefficient b_s by multiplying both sides by $\sin(2\pi st/T)$ and integrating between the same limits

We thus obtain

$$a_0 = \frac{1}{T}\int_0^T f(\tau)\,d\tau \quad . \qquad (2),$$

and for the other coefficients,

$$\left.\begin{aligned} a_s &= \frac{2}{T}\int_0^T f(\tau)\cos(2s\pi\tau/T)\,d\tau \\ b_s &= \frac{2}{T}\int_0^T f(\tau)\sin(2s\pi\tau/T)\,d\tau \end{aligned}\right\} \qquad (3),$$

where the variable has been altered for convenience in future use from t to τ.

As a_0 expresses the difference between the displacements of the point P at times $t = T$ and $t = 0$, we may, in the case of periodic motions, by choosing the time T to be very large, make a_0 as small as we like. We shall therefore neglect this quantity The remainder of the series may then be written

$$v = v_1 + v_2 + \quad v_s + \qquad (4),$$

where $$v_s = r_s \cos\{(2\pi st/T) + \theta_s\} \qquad (5),$$

r_s and θ_s being two quantities which may be determined in the usual way from a_s and b_s

Each term of the series (1) is identical in its analytical expression with what we have called a simple periodic motion giving rise to a homogeneous wave, but we must bear in mind that the equation only holds during a certain time interval, and that homogeneous light necessarily implies an infinite succession of waves Hence some care is necessary in the application of the formula. We may however, as we are at liberty to choose the time T as large as we like, express the whole disturbance as being formed by the superposition of a number of disturbances each of which may be made as nearly identical as we please with homogeneous light.

In the analytical discussion of diffraction and refraction, we have based our investigation on the treatment of homogeneous waves, and where the light was not homogeneous, we have assumed that the total effect as regards intensity, could be represented as being equal to the sum of the separate effects of a large number of homogeneous vibrations. This requires justification Imagine the disturbance, which may be of quite arbitrary type, to pass through any optical system and confine the attention to that point of the system where the observations are carried out When T is very large, we may, except possibly near the limits of time, treat each term of the series (1) as being due to a homogeneous wave, and in all cases we are taking into account, homogeneous waves are not altered in type by their passage through or reflexion from bodies Hence at the point considered, the velocity

may still be expressed as the sum of terms of the form (5), with, however, altered values of r and θ

Calculate now the average square of the velocity during the interval T The square of the right-hand side of (4) contains products such as $v_n v_s$ and hence the expression for the average value of v^2 contains terms of the form

$$\frac{r_n r_s}{T}\int_0^T \cos\{(2\pi nt/T)+\theta_n\}\cos\{(2\pi st/T)+\theta_s\}\,dt.$$

n and s being integers, the integral is easily shown to be zero The remaining terms to be considered are of the form

$$\frac{r_s^2}{T}\int_0^T \cos^2\{(2\pi nt/T)+\theta_s\}\,dt=\tfrac{1}{2}r_s^2,$$

and hence for the average value of v^2 we find $\frac{1}{2}\Sigma r^2$, but this is exactly the same expression we should have found, if we had treated each component of the series (4) as an independent homogeneous vibration. The intensity of the luminous disturbance at any time is proportional to v^2, and our proof of the independence of the separate vibrations as regards energy only applies to the average energy extended over a very long range of time The relevancy of the proposition as regards light depends on the fact that in our optical investigations we may treat the sources of light to be constant, so that the average energy is independent of the length of the time interval. This important remark was first made by Gouy*, to whom the whole of the above investigation is due. The simplification in the treatment of non-homogeneous light which was first made at the end of Art 20 now finds its complete justification, and we are at liberty, whenever it is convenient, to represent white light by superposing a number of homogeneous vibrations having periods which lie very close together But we are equally at liberty to assume any other representation so long as its resolution by Fourier's theorem gives us a distribution of intensity equal to that of the observed one. Gouy pointed out that we can regard white light as being made up of a succession of perfectly irregular impulses The type of the impulse is quite arbitrary so long as the conditions as regards distribution of intensity are satisfied

182. Application of Fourier's integral. Lord Rayleigh's investigations. Lord Rayleigh† had independently arrived at conclusions similar to those of Gouy, and has more definitely investigated the type of impulse, an aggregation of which may be considered to constitute white light‡

* *Journal de Physique*, Vol v p 354 (1886).

† *Collected Works*, Vol III p 60.

‡ *Collected Works*, Vol III p. 268

If we write (1) in the form

$$v = \sum_{s=0}^{s=\infty} \{a_s \cos(2\pi st/T) + b_s \sin(2\pi st/T)\},$$

and substitute the values of a_s and b_s from (2) and (3), we obtain

$$f(t) = \frac{2}{T}\left\{\frac{1}{2}\int_0^T f(\tau)\, d\tau + \sum_{s=1}^{s=0} \int_0^T f(\tau) \cos\{2\pi s(\tau - t)/T\}\right\} d\tau$$

If now T is allowed to increase indefinitely, we may write $\omega = 2\pi s/T$ and for the increase of ω in successive terms of the sum, $d\omega = 2\pi/T$. The first term on the right-hand side vanishes By substituting integration for summation, we obtain Fourier's theorem in the form

$$\pi f(t) = \int_0^\infty d\omega \int_0^T f(\tau) \cos \omega(\tau - t)\, d\tau \quad . \quad . \quad (6)$$

This equation represents the way in which any given function $f(t)$ may be analysed into its homogeneous components The next step is to find how much energy is to be ascribed to each small range of periods defined by the values of ω This is most easily done by means of a theorem expressed by the following equations*

$$\pi \int_0^\infty f(t)\,\phi(t)\, dt = \int_0^\infty (A_1 A_2 + B_1 B_2)\, d\omega,$$

where

$$A_1 = \int_0^{+\infty} f(\tau) \cos \omega\tau\, d\tau, \quad B_1 = \int_0^{+\infty} f(\tau) \sin \omega\tau\, d\tau,$$

$$A_2 = \int_0^{+\infty} \phi(\tau) \cos \omega\tau\, d\tau, \quad B_2 = \int_0^{+\infty} \phi(\tau) \sin \omega\tau\, d\tau.$$

If $f(t)$ expresses a vector, the square of which is proportional to the energy, $\int_{-\infty}^{+\infty} f(t)^2\, dt$ may be taken as the measure of the total energy of the disturbance, and by the above theorem,

$$\pi \int_{-\infty}^{+\infty} f(t)^2\, dt = \int_0^\infty (A^2 + B^2)\, d\omega \quad . \ldots \ .(7),$$

where

$$A = \int_{-\infty}^{+\infty} f(t) \cos \omega\tau\, d\tau, \quad B = \int_{-\infty}^{+\infty} f(\tau) \sin \omega\tau\, d\tau$$

It follows that $(A^2 + B^2)/\pi$ may be taken as the measure of the energy in the range defined by $d\omega$, the frequency being $\omega/2\pi$.

As an example, Lord Rayleigh takes a disturbance originating at a point and having at any time a velocity given by

$$f(t) = e^{-c^2 t^2} \quad . \quad \ldots\ldots \quad . \quad (8).$$

* Schuster, *Phil Mag*, Vol XXXVII p 509 (1894).

This disturbance is very small when t is large on the negative or positive side, and has a maximum for $t=0$. In this case

$$A = \int_{-\infty}^{+\infty} e^{-c^2\tau^2} \cos \omega\tau \, d\tau = \frac{\sqrt{\pi}}{c} e^{-\omega^2/4c^2},$$

$$B = \int_{-\infty}^{+\infty} e^{-c^2\tau^2} \sin \omega\tau \, d\tau = 0$$

Hence the energy included in the range between ω and $\omega + d\omega$ is

$$c^{-2} e^{-\omega^2/2c^2} d\omega$$

This represents a distribution of intensity resembling to some extent that observed in the case of the light emitted from black bodies. The example is sufficient to show that it is possible to represent white light as being due to the emission of a succession of disturbances, each of which roughly resembles that represented by (8). The larger the value of c, the more sudden will each disturbance be, approaching ultimately to an impulsive motion.

183. White light analysed by grating. We shall now have to discuss the passage of a disturbance through optical instruments. The analogy of simple sound phenomena may be of help to us in understanding the true nature of the effects produced. If a blast of air be directed against a rotating disc, perforated at regular intervals like the disc of a siren, a musical sound is heard, the continuous blast of air having a regular periodicity impressed upon it by the instrumental appliance. Similarly if a sharp impulsive motion of the air be reflected from a railing, the bars of which are at different distances from the observer, the reflected impulse returning at regular intervals of time, may produce the effect of a musical note. In both cases the regularity in the nature of the sound which is heard, has been impressed upon it by outside influence.

Taking the second example, we may also say that the musical note was really already contained in the original sound, but was mixed up with a multitude of other periods, the total effect being an impulsive motion. According to this view, the reflecting rails would simply have sifted out one period, destroying the effects of others by interference. The effect of the railing may therefore be said to be exactly analogous to that of Fourier's analysis, which by calculation picks out the simple period contained in an irregular disturbance. But from whatever point of view we look upon it, the *regularity*, which is due not so much to the presence of a simple periodic motion as to the absence of other superposed periods, is not contained in the original sound, but is produced by the regular spacing of the rails.

The effect of a grating on an impulsive motion of light is the same as that of a set of railings on an impulsive motion of sound.

In Fig. 76, Art. 59, let the incident light consist of a single impulse spread over a plane wave-front which is parallel to the grating. The impulsive motion will reach the points C_1, C_2, C_3, at regular intervals. If therefore a lens be placed in such a position that a wave-front HK would be brought together at its principal focus, a succession of impulses would pass that focus at regular intervals of time, the result being a periodic disturbance Whether we say that the grating is the cause of the periodic disturbance, or that it has only picked out a particular period already existing, the regularity in the luminous disturbance passing through the focus must be entirely ascribed to the instrumental appliance

It is interesting to follow out the effect of a grating in modifying a disturbance of any shape For this purpose we must define the action of a grating a little more closely Let s be measured along the grating, at right angles to its lines, and $f(s, t)$ be the displacement The grating modifies the disturbance in a periodic manner, and we obtain the simplest kind of modification by assuming that the disturbance in the reflected light is equal to $\cos qs \; f(s, t)$, $2\pi/q$ measuring the distance between the lines of the grating An imaginary grating having this property was made use of by Lord Rayleigh in his article on the wave theory in the *Encyclopaedia Britannica* It may conveniently be called a simple grating, and it can be shown that all real gratings may be represented by the superposition of a number of simple gratings

I have shown* that if any disturbance, originally coming from a point, is spread over a plane wave-front at right angles to x with a velocity determined by $\psi(Vt - x)$ and falls on a simple grating, the displacement in the reflected beam is determined by the equation

$$2\pi\phi(t) = \frac{h\cos\beta}{VF}\int_{-l}^{+l}\cos qs\,\psi(Vt - \gamma s)\,ds \qquad (9)$$

The displacement is measured at the focus of the lens collecting it. The other quantities which occur in the equation are defined as follows: h = height of grating, $\gamma = s(\sin\beta - \sin\alpha)$, α = angle of incidence, β = direction of reflected beam, $2l$ = width of grating

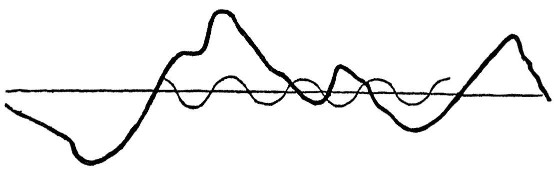

Fig 180

In Fig 180 let the thick line represent the velocity of the disturbance travelling in the positive direction, and the thin line

* *Phil Mag* Vol. xxxvii p 509 (1894)

the curve $y = \cos qx$, drawn from the point $x = 0$ to the point $x = 2\pi N/q$ Then if y' be the ordinate of the thick curve, the displacement at the focus of the telescope after reflexion from the grating at a certain time, is seen by (9) to be proportional to $\int yy'dx$ The displacement at all times is obtained by letting the wave travel forward, the cosine curve remaining in the same position. The formula (9) and the geometrical interpretation just given, bring out clearly the analogy between the action of the grating and the integration involved in the calculation of the coefficients in Fourier's series I must refer to the original paper for a detailed discussion of the application of the above equation, but the following two special cases may help the student in clearing his ideas.

Case I The incident beam is homogeneous The light reflected from the grating in any direction is then also homogeneous, and has the same period as that of the original light That is to say, whatever the periodicity of the grating, it has no power to alter the periodicity of the disturbance The distribution of amplitude in different directions is the same as that which has already been obtained in the Chapter on Gratings

Case II The incident disturbance consists of a single impulsive velocity The disturbance at the focus of the telescope consists then of an impulsive displacement followed by a vibration represented by a cosine curve, continuing for as many periods as there are lines in the grating. The period is the same as that of the homogeneous vibration which, having the wave-front parallel to that of the disturbance considered, would have its principal maximum at the focus of the telescope, that is to say, the periodicity in the reflected beam depends now on the direction in which the telescope points

184. White light analysed by dispersive media. The mechanism by means of which a grating converts an impulsive motion into a regular succession, with one predominant period, is easily explained The action of a prism is a little more difficult to understand Nevertheless, we know that a prism behaves in the same way as a grating, and the method of its action suggests itself as soon as we consider that the separating power of the prism is due to the unequal velocity of different wave-lengths through its substance. An impulsive motion therefore, started in a dispersive medium, cannot remain an impulsive motion, but the disturbance lengthens out as the wave proceeds Having again recourse to Fourier's theorem, we may analyse an impulse, and imagine it to be made up of a number of different groups of waves of different lengths. These groups, according to Art 180, are propagated with different velocities so that a separation of the different wave-lengths takes place, except in the

particular case considered in that article, in which the law of propagation is such that the group velocity is the same for all wave-lengths This being the simplest conceivable law, we may consider the action of a prism made of a substance for which the group velocity is constant. A plane impulsive motion reaching such a prism obliquely is refracted, and we can draw a plane over which the disturbance is spread in the prism in the ordinary way, by substituting the group velocity for the wave velocity. The plane of the disturbance stands therefore oblique to what, in ordinary refraction, would be the wave-front, and the optical distance from the original wave to the different points of the plane of disturbance is not the same.

It has also been pointed out that as the group proceeds, maxima are periodically changed into minima, and vice versa On the plane of disturbance therefore the motion will not be everywhere in the same direction, but will change alternately from one direction to the other If we now follow this plane of disturbance proceeding with the group velocity as it is refracted out of the prism, we obtain another plane of disturbance oblique again to what, under ordinary circumstances, would be the wave-front. Its position would be the same as that of an ordinary wave-front which has passed through the prism with the group velocity If the emergent wave be now received by a lens, the disturbance at the focus of the lens consists of a periodic motion, the different parts of the plane of the disturbance passing through the focus at different times The greater the resolving power of the prism, the greater will be the number of inversions in the plane of disturbance after it has left the prism and therefore the more will the light passing through the focus of the lens be homogeneous In this way we may convince ourselves that the action of a prism is identical with that of a grating Although our reasoning is strictly correct only for a prism composed of a material which has a definite law of dispersion, the result must be the same in other cases, because, fixing our attention on a certain narrow range of wave-lengths, we may always consider the groups of waves within this range to proceed through the material with constant velocity We may therefore apply the above reasoning to each such narrow range separately

185. Interference. In Chapter IV we based our investigations on interference on the consideration of homogeneous waves, and in the case of white light we imagined a large number of such homogeneous waves covering all the wave-lengths given out by the luminous source This method of procedure is perfectly correct, and must always lead to the right results, but if we are right in saying that we may consider white light as being due to a short impulsive motion we should also be able to obtain accurate results by considering a single

impulse, in place of a superposition of homogeneous vibrations. In most cases the distribution of intensity along the spectrum does not affect the interference phenomena, and we may therefore assume for convenience any arbitrary form of the impulse, such as an instantaneous velocity imparted at a point.

It is very instructive to follow out this view in a few simple cases. What happens in ordinary interference experiments when the original disturbance consists of simple impulses? Take, for instance, two slightly inclined mirrors and the ordinary arrangement of Fresnel's mirror experiment. At any point at which one mirror would give us a single impulse, we now have two impulses succeeding each other after an interval of time which depends on the point of observation. Looked at from our present standpoint, the word interference seems quite inappropriate to this case; for how can two separate motions following each other be said to interfere? There is indeed no true interference here. But let now the two impulses succeeding each other enter a spectroscope of high resolving power. If the dispersion is produced by a grating, we have seen how each disturbance is converted into a regular succession of waves, and the waves due to each of the two impulses are now made to overlap. According to the position of the observing telescope the phases may coincide or be in disagreement, so that there may be a weakening or strengthening of the resulting intensity. The interference therefore only comes into play when the impulse is spread out into waves by the instrumental appliance. The instrument must first make more or less homogeneous light before interference can take place.

When Fresnel's mirrors are used, in conjunction with white light and without any spectroscopic help to separate the wave-lengths, we yet observe a certain periodicity in the coloration indicating interference, and this may appear to the student to be a difficulty in accepting the statement that what he sees may simply be explained by a succession of two impulses, which can only interfere after having been spread out by some optical arrangement. But in this case, the apparent interference is really due to a peculiarity in our physiological sensations which here act similarly to the presence of a grating in sifting out certain periodic motions. Imagine two blows directed against the weight of a pendulum. The resulting motion depends on the interval of time which elapses between the blows. If this is equal to an exact multiple of the complete period of vibration of the pendulum, the effects of the blows will assist each other to produce a greater velocity. On the other hand, an interval of half the time of vibration would cause the second blow to neutralize the effect of the first. Similarly when two luminous impulses enter the eye, we must imagine them to affect our nerves, which are tuned to the three

primary colours The interference therefore here is in the eye, and is not objective. If the interference effects due to Fresnel's mirrors are observed with a bolometer, they are much less marked, showing only one minimum and one maximum, and this residual periodicity is due to the fact that in the white light which is at our command the different periods are not equally represented, but a pronounced maximum of intensity appears at a certain region of the spectrum The light therefore begins to approach some kind of regularity which gives results approaching those observed with homogeneous light I must refer the reader to a paper on Interference Phenomena*, in which these matters have been discussed in greater detail

Considerable discussion took place at one time as to the determination of a supposed property of light called its "regularity" apart from its "homogeneity" The power of interfering with a long difference of path was used as a test of this so-called regularity Hence importance was attached to certain experiments in which even with white light interference effects were produced when the retardation amounted to as much as 50,000 or 100,000 wave-lengths Gouy and Rayleigh however pointed out that in these cases the limit of the observed interference effects is entirely determined by the resolving instrument, and not by the light This follows directly from the consideration of the action of gratings and prisms which has been given in the preceding articles As this matter was, and is still, often badly understood, we may perhaps trace the manner in which the misconception has arisen. It was found that light spoken of roughly as being homogeneous, showed interference effects only up to certain differences of path, and that two different sources of light giving out such so-called homogeneous vibrations, could not interfere with each other The explanation of this fact was based on the idea that a molecule sends out a succession of regular vibrations during a certain period, at the end of which it is disturbed by an impact, or otherwise Light sent out from such a molecule was described as being homogeneous light, with the peculiarity that the phase of the motion was subject to sudden changes at more or less regular intervals of time The fundamental error here is, that such light should be described as being homogeneous. It has been pointed out on several occasions in these pages that a homogeneous vibration implies an infinite succession of waves of the same length, and without any such peculiarity as a change of phase would imply We have shown that the action of a grating is equivalent to that of Fourier's analysis, and anyone can convince himself that a succession of sine or cosine curves joined together so as to represent certain changes of phase can only be reproduced by means of a Fourier's

* *Phil. Mag* Vol. xxxvii p 509 (1894)

series including a large number of different wave-lengths. It may be perfectly correct, that as regards the mechanism causing the limits of interference the above explanation is true, and we should in that case be justified in saying that the regularity of the motion is interfered with by molecular impacts, but this is the same thing as saying that the homogeneity of the light sent out is destroyed by these impacts. Without forming any hypothesis as to the mechanical cause which produces the irregularity, we simply state the facts if we ascribe the limit of the power of interference to the want of homogeneity of the source The reason for the absence of homogeneity may then be left over for an independent discussion. Experimental investigation of the retardation at which interference is still possible coincides therefore with the investigation as to the homogeneity of the light Its importance begins when we come to study the details of the structure of the radiations of luminous gases which in ordinary language are described as being homogeneous.

The explanation which depends on the assumption of sudden changes of phase by molecular impacts, is also generally used to account for the absence of interference of two different sources, and here again the mechanical explanation may be perfectly correct. But it is important to realize that two independent sources sending out homogeneous light of exactly the same wave-length are capable of interfering with each other in exactly the same way as if the light were derived from the same source. If two waves, spreading out from *e g.* two mercury lamps, which give out nearly homogeneous radiations are found not to interfere with each other, it only means that the homogeneity is not sufficient.

When an impulsive motion of definite type is analysed by Fourier's series, it is found that there is a certain definite relation of phase between the waves of different periods. Some explanation may be necessary to bring this into agreement with our previous view of white light, according to which there could be no relation of phase between any two wave-lengths however near each other. It is obvious at once that no phase relation can exist when the light is such that its average intensity is constant. In fact, such a relation is inconceivable. As it exists for a single impulse, it must be destroyed by the succession of impulses which converts the instantaneous source of light into one of constant average intensity

186. Talbot's Bands. If the spectrum formed by a prism or grating is observed, half the pupil of the eye being covered with a thin plate of mica or glass, the spectrum is seen to be traversed by dark bands, provided the plate be inserted on that side on which the blue of the spectrum appears. These bands were first observed by

Fox Talbot. Instead of viewing the spectrum directly we may use a telescope, the plate being inserted on the side of the thin edge of the prism forming the spectrum, so as to cover a portion of the aperture of the object glass.

Similar bands have been observed by Powell, who used a hollow glass prism with its refracting edge pointing downwards and filled with some highly refractive liquid, into which he inserted a plate of glass with its lower edge parallel to the edge of the prism and so that its plane approximately bisected the angle of the prism The plate was only partially inserted, so as to leave the lower portion clear The bands only appeared when the refractive index of the liquid was greater than that of the glass, but Stokes showed that when the refractive index of the glass was the greater of the two the band could still be observed, only in this case it was necessary to place the plate in the thinner part of the prism, leaving the thicker portion clear

A simple explanation of these bands is sometimes based on the consideration that the two portions of the light, which, in the absence of the interposed plate, would reach the retina in the same phase, are retarded relatively to each other by the plate, so that interference may take place This explanation is obviously incomplete, for it leaves out of account the essential fact that the effects are only observed when the plate is inserted on one side and not on the other. A more complete explanation taking account of this want of symmetry has been given by Airy and Stokes, and involves an elaborate mathematical process A very simple treatment may be given if, instead of basing the calculation on Fourier's analysis, we consider the source of the light to be due to a succession of impulsive velocities In Fig 76 (Art 59) we have a wave-front consisting of a simple impulse which reaches the grating so that the points A_1, A_2, A_3, etc. are simultaneously disturbed At the plane HK, the disturbance will reach the points C_1, C_2, C_3, in succession, and if a lens be placed with its axis at right angles to HK, the disturbance will pass the focus of the lens at regular intervals of time, as already explained in Art 183

The question now is How can the impulses which succeed each other at the focus of the lens, be made to interfere with each other? Clearly only by retarding those which reach the focus first or by accelerating those which reach it last A plate of appropriate thickness introduced on the left-hand side of the figure as it is drawn could be made to answer the purpose If, on the contrary, the same plate be introduced on the right-hand side, it would only retard those impulses which already arrive late, and therefore no interference could take place This is all that need be said in explanation of the bands, but a more

detailed consideration leads to a simple expression for the calculation of the thickness of the plate which shows the bands most distinctly.

It is easily seen that the best thickness is secured when the whole series of impulses is divided into two equal portions, the impulses arriving in pairs simultaneously at the focus. The retardation must therefore be such that the retarded impulse coming from the first line of the grating, and the unretarded impulse coming from the central line, arrive together. This means that the retardation is $\frac{1}{2}N\lambda$, if N is the total number of lines in the grating, and the plate should be pushed sufficiently far into the beam to affect half its width. The wave-length λ here means the wave-length of that homogeneous train of waves which has its first principal maximum at the focus of the lens, so that the retardation of each impulse compared with the next is λ. If the retardation is either greater or smaller, some of the impulses arrive too soon or too late to overlap others, and the bands are less clear. If the retardation has more than twice its most effective value, the series of impulses from the first half of the grating pass through the focus later than those coming from the second half, and hence there cannot be any interference.

If at a certain point of the spectrum corresponding to a wave-length λ there is a maximum of light, the relative retardation of the two interfering impulses must be equal to $m\lambda$, m being an integer; the next adjoining band towards the violet will appear at a wave-length λ' such that $m\lambda = (m+1)\lambda'$.

Hence for the distance between the bands

$$\frac{\lambda-\lambda'}{\lambda'} = \frac{1}{m},$$

with the best thickness of interposed plate, $m = \frac{1}{2}N$, and hence $(\lambda-\lambda')/\lambda' = 2/N$ where λ' in the denominator may with sufficient accuracy be replaced by λ. If λ'' be that wave-length nearest to λ at which there is a minimum of light, it follows that

$$\frac{\lambda-\lambda''}{\lambda} = \frac{1}{N}.$$

If a linear homogeneous source of light of wave-length λ be examined by means of a grating, the central image extends to a wave-length λ_1 such that

$$\frac{\lambda-\lambda_1}{\lambda} = \frac{1}{N},$$

where N, as before, is the total number of lines on the grating.

Hence the following proposition:—If, in observing Talbot's bands, the best thickness of retarding plate be chosen, the distance between each maximum and the nearest minimum is equal to the distance between

the central maximum and the first minimum of the diffractive image of homogeneous light, observed in the same region of the spectrum with the same optical arrangement. This proposition holds for all orders of spectra, but the appropriate thickness of the retarding plate increases in the same proportion as the order.

If we use prisms instead of a grating, the number of lines N must be replaced by the quantity which corresponds to it as regards resolving power, viz., $td\mu/d\lambda$ where t is the aggregate effective thickness of the prisms. It follows that the retardation which gives the best interference bands with prisms, is $\frac{1}{2}\lambda t d\mu/d\lambda$. The above explanation was given, with further details, in a recent communication to the *Philosophical Magazine**.

187. Roentgen radiation. The radiations which are produced by the impact of kathode rays on solid bodies, possess, as shown by their discoverer, some properties which at first seem to distinguish them from the transverse waves of light. Roentgen could find no trace of refraction or polarization, and he tentatively suggested that the rays he had discovered were due to longitudinal vibrations. The absence of diffraction and interference was also felt to be a difficulty in ascribing the observed phenomena to a wave-motion. It was however almost immediately pointed out† that according to Sellmeyer's equation waves of very short length would not suffer any refraction, the velocity of propagation through all media being the same, and that the want of homogeneity would be quite sufficient to account for the apparent absence of interference. The absence of polarization by reflexion or refraction follows from the equality of the velocities of propagation. Some years later Professor J. J. Thomson‡ calculated the electromagnetic disturbance sent out by the impulsive stoppage of an electrified particle moving with great velocity. Such an electromagnetic wave possesses all the properties we have ascribed to white light, and would only be distinguishable from it by the smallness of the linear quantity involved.

Thus in the hypothetical law of radiation expressed by (8), the disturbance is the more sudden the greater the value of c, and the linear quantity, V/c, is a measure of the distance in space over which the disturbance is appreciable. If this view as to the nature of the Roentgen rays is right, it would still be correct to say that these rays are due to very short waves, the spectrum being continuous and extending over a considerable range of wave-lengths. If there is any distinction between the statements that Roentgen rays consist of short

* *Phil. Mag.* Vol. VII. p. 1 (1904).
† *Nature*, Vol. LIII. p. 268 (1896).
‡ *Phil. Mag.* Vol. XLV. p. 172 (1898).

waves and that they are due to irregular impulses, it can only lie in the fact that the first statement excludes long waves and the second does not

The matter is introduced here because the mechanical stoppage of a moving electrified particle may serve as a good example of the kind of wave we imagine white light to consist of

188. The radiation of a black body. All bodies of sufficient thickness send out radiations which are independent of the nature of the body, and therefore identical with those sent out by a perfectly black surface. The experimental investigation of the law of complete radiation of a black body presents considerable difficulties which have only recently been partially overcome.

According to Stefan's law the intensity of the total radiation is proportional to the fourth power of the absolute temperature. This law, for which there is some theoretical justification, has been found to be correct so far as our observations are able to go We cannot enter here into a complete account of the subject, but must refer to the papers of W. Wien*, and others, whose work has been summarized in an important paper by H Rubens and F. Kurlbaum†.

The distribution of energy in the spectrum seems to be best represented by a formula first given by M Planck If C is a constant, and the energy lying between the wave-lengths λ and $\lambda + d\lambda$ be $Ed\lambda$,

then $$E = C\lambda^{-5}/(e^{\frac{c}{\lambda T}} - 1).$$

If this formula is examined it will be found that for a given value of the absolute temperature T the value E is a maximum, and that this maximum satisfies the condition $\lambda_{\max} T = \text{const}$ If the quantity E is calculated for that value of λ at which it is a maximum it is found to be proportional to the fifth power of the absolute temperature These two important relations were derived by W. Wien from theoretical considerations

The manner in which the complete radiation of a body is independent of its nature, and yet vibrations of individual molecules as observed by spectroscopic analysis are entirely characteristic of the nature of the body, forms a great difficulty, in the satisfactory treatment of the subject of radiation

Attempts have been made to show that the white light of the complete radiation is an entirely distinct phenomenon from the homogeneous vibration which we observe in the spectra of gases But the theory of exchanges shows that there must be some connexion

* *Wied Ann.* Vol LVIII p. 62 (1896).

† *Ann. d Phys.* Vol. IV p 649 (1900).

between the two It is, however, quite possible to imagine causes for the radiation of white light in addition to the molecular forces which cause homogeneous light Thus we note that at any rate in conducting bodies there are free electrons capable of independent motion through the substance, and these electrons, which we may imagine to behave something after the nature of a gas, would send out radiations, whenever their velocity is altered in magnitude or direction The radiation sent out by such moving electrons would therefore be exactly of the same kind as that calculated by J J Thomson for the explanation of the Roentgen radiations.

189. Doppler's Principle. The investigation of light sent out by moving bodies, or passing through moving bodies, lies beyond the scope of this book, but in considering the radiation sent out by gaseous molecules it is impossible to ignore altogether the effect of motion It may be proved by elementary considerations that if the time of vibration of a moving particle be τ_0, an observer standing in the line of motion will receive vibrations, the periodic time (τ_1) of which is given by

$$\tau_1 = \frac{V \mp v}{V} \tau_0,$$

where V and v are the velocities of light and of the moving body respectively and the minus or plus sign holds according as the particle approaches or recedes. The effect being, at any rate as regards the first order of magnitude, one of relative motion, the equation also holds if the particle is stationary and the observer moves

190. Homogeneous radiations It is a matter of considerable interest to enquire how far the vibrations of gaseous molecules are observed to be homogeneous Although each molecule considered as stationary may continuously perform simple vibrations, the considerations explained in the last article show that owing to the motion of the molecules there must be a natural limit to the sharpness of the line.

This was first pointed out by Lippich, and later by Lord Rayleigh, who calculated that at the temperature of a Bunsen burner, light cannot be reduced to narrower limits than about the one-hundredth part of the interval between the sodium lines In a further paper Lord Rayleigh* discusses the effect of this widening of the lines on the visibility of interference bands with large differences of path. We shall here content ourselves with giving the results of some experimental investigations

We may obtain the highest degree of homogeneity by observing the spectra under reduced pressures, and under these circumstances we

* *Coll Works*, Vol. III. p 258

are unable to tell what velocities to ascribe to the molecules. It is therefore difficult to say how much of the apparent want of homogeneity is really due to motion. Michelson* has deduced from the visibility curve obtained by observation with his well-known instrument the intimate structure of a number of lines of Cadmium, Mercury, and Thallium. The observations on interference for large differences of path have been pushed still further by Fabry and Perot†. The limits of interference of the green Cadmium lines were arrived at with a retardation of 472,000 wave-lengths, and this number was still exceeded with the green Mercury line, where the limits of interference were only reached with a difference of path of 790,000 wave-lengths.

Lummer and Gehrcke‡ by a method similar in principle to that of Fabry and Perot were able to observe the intimate structure of some of the Mercury lines, and found that even the lines which in spectroscopes of considerable resolving power appear as homogeneous lines, split up into a number of components. In the case of the Mercury line having wave-lengths $5{\cdot}46 \times 10^{-5}$, twenty-one components were counted.

The careful investigation of homogeneous radiation is of some importance, as it enables us to form an estimate of the degree to which different atoms of the same substance are identical. Clerk Maxwell in an eloquent passage of his lecture on "Molecules," delivered to the British Association at Bradford, has pointed out the value of spectroscopic investigation in showing that the atoms of hydrogen in our laboratories have the same properties as those found in distant regions of stellar space. Maxwell based his conclusions on the fact that by spectroscopic measurements we can identify the wave-length to one part in ten thousand. Since the time that lecture was delivered the accuracy of measurement has been increased many times. The most recent determinations of Lummer and Gehrcke do not lend themselves very easily to a mathematical estimate, and I therefore fall back on the conclusions derived by Michelson from his visibility curves. I take the green line of thallium, and find that if $\delta\lambda$ represents the difference in wave-length between the point at which the intensity of the line is a maximum, and that at which it has fallen to half its maximum value, $\delta\lambda/\lambda = 478 \times 10^{-9}$. Forming an estimate from the general shape of the intensity curve, we may judge from it that about one-quarter of the total intensity of the line is due to light differing from the central wave-length by more than $\delta\lambda$. The greater part of the width of the

* *Phil. Mag.*, Vol. XXXIV. p. 280 (1892).

† *C. R.* Vol. CXXVIII. p. 1223 (1899).

‡ *Ann. d. Physik*, Vol. X. p. 457 (1903).

line is shown by Michelson to be due to molecular motion, but I will assume that different molecules are not really exactly like each other, but differ slightly in their period of vibration from some average value As an extreme case, I will imagine that half of the light extending beyond the limit $\delta\lambda$ is due to such a real difference in period It would follow that about 12 % of the molecules would show a real difference in wave-length of more than 478×10^{-9}, which is approximately one part in two millions Let us see for a moment what this means

If of two tuning-forks, vibrating about one hundred times a second, one gains a vibration on the other in 1000 seconds, so that the periods differ by one part in one hundred thousand, it is just possible by accurate measurement to detect this difference, but it requires several days' work to make the comparison with sufficient accuracy. No maker would, however, undertake to supply you with ten tuning-forks with a guarantee that not more than one of them would give a note differing from that of the average by one vibration in one hundred thousand, and we have seen that the atoms, as regards their vibration, must be at least twenty times as well in agreement as this

If you had a great many clocks, and found that, taking their average rate to be correct, not more than one in eight would be wrong by a second in twenty-three days, that would represent the maximum amount of variation which our interpretation of the experiment allows us to admit in the case of molecular vibrations But would any maker undertake to supply you with a number of clocks satisfying that test? If further, it is considered that the limits we have chosen for the possible variations of molecular vibrations are far too great, we see that though Sir John Herschel's saying, that atoms possess the essential character of manufactured articles, is still correct, yet no manufactured article approaches in accuracy of execution the exactitude of atomic construction. We may conclude with Maxwell that: "Each molecule therefore, throughout the universe, bears impressed on it the stamp of a metric system as distinctly as does the metre of the Archives at Paris, or the double royal cubit of the temple at Karnac."

INDEX.

The numbers indicate reference to pages.

CAMBRIDGE PRINTED BY J AND C F CLAY AT THE UNIVERSITY PRESS

A TEXT-BOOK OF PHYSICAL CHEMISTRY. By Dr R. A LEH FELDT, Professor of Physics at the East London Technical College With 40 Illustrations Crown 8vo, cloth, 7s 6d

THE CHEMICAL SYNTHESIS OF VITAL PRODUCTS AND THE INTER-RELATIONS BETWEEN ORGANIC COMPOUNDS. By Professor RAPHAEL MELDOLA, F R S, of the City and Guilds of London Technical College, Finsbury Two vols Vol I now ready Super Royal 8vo, 21s net

A TEXT-BOOK OF PHYSICS. With Sections on the Applications of Physics to Physiology and Medicine By Dr R A LEHFELDT Fully Illustrated Crown 8vo, cloth, 6s

PHYSICAL CHEMISTRY FOR BEGINNERS. By Dr Ch. M VAN DEVENTER With a Preface by J H VAN 'T HOFF Translated by Dr R. A LEHFELDT, Professor of Physics at the East London Technical College 2s 6d

A FIRST YEAR'S COURSE OF EXPERIMENTAL WORK IN CHEMISTRY By ERNEST H COOK, D Sc., F I C, Principal of the Clifton Laboratory, Bristol With 26 Illustrations Crown 8vo, cloth, 1s 6d

AN EXPERIMENTAL COURSE OF CHEMISTRY FOR AGRICULTURAL STUDENTS. By T S DYMOND, F I C, Lecturer on Agricultural Chemistry in the County Technical Laboratories, Chelmsford With 50 Illustrations 192 pages Crown 8vo, cloth, 2s 6d

THE STANDARD COURSE OF ELEMENTARY CHEMISTRY. By E J COX, F C S With 90 Illustrations 350 pages Crown 8vo, cloth, 3s Also obtainable in five parts, limp cloth Parts I-IV, 7d each, Part V, 1s

A PRELIMINARY COURSE OF PRACTICAL PHYSICS. By C E ASHFORD, M A, Headmaster of the Royal Naval College, Osborne Fcap 4to, 1s 6d

PHYSICAL DETERMINATIONS. A Manual of Laboratory Instructions for the Determination of Physical Quantities By W. R KELSEY, B Sc., A I E E, Lecturer in Physics to the Bradford Municipal Technical College 4s 6d

ELECTROLYTIC PREPARATIONS. Exercises for use in the laboratory by chemists and electro-chemists By Dr KARL ELBS, Professor of Chemistry at the University of Giessen Translated by R S HUTTON, M Sc Demy 8vo, 4s 6d net

THE ELECTRIC FURNACE. By HENRI MOISSAN, Professor of Chemistry at the Sorbonne Authorized English Edition Translated by A T DE MOUILPIED, M Sc, Ph D With Illustrations. Demy 8vo, 10s 6d net

ELECTRICAL TRACTION. By ERNEST WILSON, Wh.Sc., M.I.E.E. Professor of Electrical Engineering in the Siemens Laboratory, King's College London Crown 8vo, 5s

ELECTRICITY AND MAGNETISM. By C. E ASHFORD, M A. Headmaster of the Osborne Royal Naval College, late Senior Science Master at Harrow School With over 200 Diagrams Crown 8vo, 3s 6d.

MAGNETISM AND ELECTRICITY. An Elementary Treatise for Junior Students, Descriptive and Experimental By J PALEY YORKE, of the Northern Polytechnic Institute, London With nearly 150 Illustrations 3s 6d

THE EVOLUTION THEORY. By AUGUST WEISMANN, Professor of Zoology in the University of Freiburg im-Breisgau Translated by Professor J ARTHUR THOMSON With numerous Illustrations and Coloured Plates Two Vols Royal 8vo, 32s net

The importance of this work is twofold In the first place, it sums up the teaching of one of Darwin's greatest successors, who has been for many years a leader in biological progress As Professor Weismann has from time to time during the last quarter of a century frankly altered some of his positions, this deliberate summing up of his mature conclusions is very valuable In the second place, as the volumes discuss all the chief problems of organic evolution, they form a reliable guide to the whole subject, and may be regarded as furnishing—what is much needed—a Text-book of Evolution Theory.

ANIMAL BEHAVIOUR. By C. LLOYD MORGAN, LL D, F.R.S., Principal of University College Bristol, author of 'Animal Life and Intelligence,' etc With numerous Illustrations Large crown 8vo, 10s. 6d

HABIT AND INSTINCT. By C LLOYD MORGAN, LL D., F R.S. With Photogravure Frontispiece viii+352 pages Demy 8vo, cloth, 16s

A TEXT-BOOK OF ZOOLOGY. By G P. MUDGE, A.R.C.Sc. Lond, Lecturer on Biology at the London School of Medicine for Women, and the Polytechnic Institute, Regent Street With about 200 original Illustrations Crown 8vo, cloth, 7s 6d

ELEMENTARY NATURAL PHILOSOPHY. By ALFRED EARL, M A, Assistant Master at Tonbridge School With numerous Illustrations and Diagrams Crown 8vo, cloth, 4s 6d

A CLASS-BOOK OF BOTANY. By G P. MUDGE, A.R.C.Sc. Lond., F Z S, and A J MASLEN, F L S, Lecturer on Botany at the Woolwich Polytechnic With over 200 Illustrations Crown 8vo, 7s 6d

THE BECQUEREL RAYS AND THE PROPERTIES OF RADIUM. By the Hon. R J STRUTT, Fellow of Trinity College, Cambridge. With Diagrams Demy 8vo 8s 6d net

A MANUAL OF ALCOHOLIC FERMENTATION AND THE ALLIED INDUSTRIES. By CHARLES G MATTHEWS, F.I C., F.C S., etc Fully Illustrated Crown 8vo, cloth, 7s 6d. net

WOOD. A Manual of the Natural History and Industrial Applications of the Timbers of Commerce By G. S BOULGER, F L.S, F.G S. Fully Illustrated Crown 8vo, 7s 6d net

PSYCHOLOGY FOR TEACHERS. By C LLOYD MORGAN, LL D, F R S xii+251 pages Crown 8vo., 3s. 6d

ANIMAL SKETCHES. By C. LLOYD MORGAN, LL.D, F.R.S. viii+312 pages, with 52 Illustrations (many of them full-page) Crown 8vo., cloth, 3s 6d